Economic Theory
https://doi.org/10.1007/s00199-026-01740-1

RESEARCH ARTICLE

Trade policy in a model of sequential production

Hartmut Egger[1,2] · **Christian Fischer-Thöne**[3]

Received: 14 May 2025 / Accepted: 14 July 2026
© The Author(s) 2026

Abstract
We develop a two-country trade model with sequential production in which a technologically backward economy exports intermediates in exchange for final goods from a technologically advanced economy. The advanced country can raise welfare by imposing import tariffs on intermediates, whereas export tariffs on intermediates or import tariffs on final goods are ineffective for the backward economy. Instead, the backward economy resorts to a quantitative export constraint, diverting part of its labor force into a less efficient autarkic production process. We show that non-cooperative policy choices reduce trade and lower welfare in both countries, underscoring the need for a trade agreement. To sustain mutually beneficial liberalization after prior tariff reductions, however, an agreement may have to link tariff cuts to the removal of export constraints, particularly when lifting these constraints is costly for the backward economy.

Keywords Sequential production · Global value chains · Trade policy · Deep trade agreements

JEL Classification F12 · F13 · F68

1 Introduction

Global value chains have expanded substantially over the past decades and now account for a large share of international trade (see Johnson and Noguera 2017; Johnson 2018). Their growing importance has stimulated a rich theoretical literature aiming to understand how production is organized across countries (e.g., Antràs and Chor 2013; Costinot et al. 2013; Antràs and de Gortari 2020). A central insight of this work is that the *sequence* of production stages matters for both the allocation of tasks and the distri-

✉ Hartmut Egger
hartmut.egger@uni-bayreuth.de

1 University of Bayreuth, CESifo, GEP and IfW, Bayreuth, Germany

2 Department of Economics, University of Bayreuth, Universitaetsstr. 30, 95447 Bayreuth, Germany

3 University of Bayreuth, Bayreuth, Germany

Published online: 01 August 2026

bution of production rents.[1] Models of sequential production emphasize the exchange of intermediates for final goods as a defining characteristic of global value chains, with important implications for how trade policies operate along the value chain. This distinction is particularly important because policy can affect a value chain through two conceptually different margins: it can change the price at which an intermediate input is traded, or it can change the quantity of the input that remains available for downstream processing. In a sequential technology with indispensable stages and strong input complementarities, the quantity margin is especially relevant: restricting the physical flow of intermediates can affect the allocation of production stages along the value chain. Indeed, the vulnerability of multi-stage production to policy intervention has been highlighted in both academic and policy debates. Yet, despite this relevance, we still know relatively little about the implications of trade policy in production-stage frameworks such as Costinot et al. (2013). Existing contributions, including Antràs and Staiger (2012a), Antràs et al. (2024), Grossman et al. (2024), and Blanchard et al. (2026), represent important steps, but a comprehensive analysis of the effectiveness of tariff- versus quantity-based instruments along the value chain is still missing.

In this paper, we address two underexplored questions in the trade policy literature on global value chains: *(i)* Are conventional tariff-based instruments equally effective when trade involves the exchange of intermediates for final goods, or do quantity-based instruments become more powerful in such settings? *(ii)* Are provisions disciplining the use of quantitative export restrictions a necessary component of trade agreements that aim to secure free trade along the global value chain? While tariffs are the central policy instrument in standard trade models, their effectiveness in global value chain environments is far less well understood. Recent contributions by Antràs and Staiger (2012a) and Antràs and Staiger (2012b) show that tariff-based instruments may fail to restore efficiency when international transactions are not fully governed by market-clearing conditions, for instance, when bargaining frictions affect surplus division along the value chain (see Antràs and Chor 2022, for discussion). Our analysis complements this line of work by demonstrating that tariff ineffectiveness can arise even in the absence of such frictions: in a perfectly competitive setting, the combination of sequential production and strong input complementarities alone is sufficient to undermine the effectiveness of tariff-based instruments for technologically backward economies.

To study the effectiveness of tariff- and quantity-based trade policy instruments in global value chains, we build on the sequential production model of Costinot et al. (2013), in which the production of a single final good proceeds through multiple stages carried out in different countries. Each stage combines the intermediate input from the previous stage with local labor, and countries are positioned along the value chain according to their technological capabilities, with the most advanced economy performing the final stage. Focusing on two countries, we analyze how each can use trade policy instruments to appropriate production rents. We show that the technologically advanced economy can raise welfare by imposing an ad-valorem tariff on imported

[1] Baldwin and Venables (2013) distinguish between sequential "snakes" and non-sequential "spiders" as alternative structures of global value chains.

intermediates, whereas import tariffs on final goods or export tariffs on intermediates are ineffective for the technologically backward economy, motivating our focus on quantity-based instruments as an alternative strategic tool.

The asymmetry in the effectiveness of tariff-based instruments is a direct consequence of two features of the model proposed by Costinot et al. (2013): *(i)* a predetermined sequence of indispensable production stages, and *(ii)* strong complementarity at each stage due to the Leontief technology, which prevents substituting a lower intermediate input with a larger labor input. Under these assumptions, an import tariff imposed by the technologically advanced country becomes a non-distortionary instrument of rent appropriation: it does not alter the allocation of production stages and leaves the total output of the consumption good unchanged. By contrast, tariff-based instruments are ineffective for the technologically backward economy. They merely redistribute rents between domestic workers and the local government, without affecting the production structure or the terms of trade with the advanced economy. Thus, to counter the rent appropriation by its trading partner further down the value chain, the backward country must rely on a different, quantity-based instrument.

One such instrument that we examine in greater detail is an export constraint implemented in the technologically backward country by limiting worker mobility into the export sector and channeling labor into a less efficient autarkic production technology. Although such quantitative export constraints are generally prohibited under WTO rules (with only a few exceptions), they remain relevant trade policy instruments, particularly in developing countries. Historically, similar quantity-based measures were widely used by export marketing boards in agriculture and by state-controlled extractive sectors, both of which restricted export volumes as part of their policy mandates. These practices illustrate a long-standing tendency of upstream commodity exporters to rely on quantity-based instruments rather than price-based tariffs. More recent data confirm the continued relevance of such measures: the OECD has systematically collected information on export restrictions for more than fifteen years, and OECD (2025) document that quantitative measures — such as export quotas and export prohibitions — have gained prominence over the last fifteen years.[2] This pattern may help explain why many recent trade agreements include explicit provisions disciplining the use of quantitative export restrictions, despite their formal exclusion under the WTO framework. For example, in a systematic review of trade agreements notified to the WTO, Wu (2020) finds that "*more than three-quarter [of them] have some form of legal discipline concerning export restrictions*" (p. 97).[3]

Against this empirical background, we examine how tariff- and quantity-based instruments shape the incentives and strategic interaction of the two countries along the value chain. Within the sequential production framework described above, we

[2] Fliess et al. (2014, p. 30) report that in 2012 *[t]he share of world exports subject to quantitative export restrictions was highest for antimony (50% of world exports), tungsten (45%), magnesite (44%), fire-clay (43%), and rare earths (39%)*."

[3] Beyond the role of export quantity constraints, recent work has documented the breadth of policy areas covered by deep trade agreements. Laget et al. (2020) identify 18 core provisions commonly found in such agreements, including disciplines on technical barriers to trade, anti-dumping and countervailing measures, investment rules, competition policy, public procurement, and intellectual property rights. Hofmann et al. (2019) provide an extensive overview of these provisions.

analyze a non-cooperative trade policy game in which each country seeks to appropriate production rents. A key finding is that, under non-cooperation, these incentives reduce trade in the open economy and generate welfare losses. The technologically advanced country has an incentive to set its import tariff at the highest level consistent with the incentive compatibility of foreign workers to be employed in global value chain production, given their outside option in the autarkic technology. A higher import tariff reduces the rents accruing to the technologically backward economy and thereby strengthens its incentive to restrict intermediate exports through a quantitative constraint. Such an export constraint makes the backward country a less attractive production location, shifting additional stages of the value chain toward the advanced economy. If the backward economy ultimately ceases exporting intermediates, production in both countries reverts to the autarkic technology, eliminating global value chain production and the associated gains from trade.

This non-cooperative outcome provides the benchmark for our analysis of trade agreements. We proceed in two steps. First, we ask whether an agreement can restore the efficient open-economy allocation when countries move from the unconstrained non-cooperative equilibrium to free trade. Second, we consider the more institutionally relevant environment in which previous multilateral liberalization has already reduced tariff policy space to different degrees. We interpret such pre-existing tariff commitments as the legacy of multilateral liberalization under the WTO. The key question is then whether, after tariffs have already been partly liberalized, countries can still reach a mutually beneficial agreement along the value chain, and whether such an agreement can remain shallow or must instead be deep by also disciplining quantitative export restrictions.

Cooperation between the two countries can help overcome the prisoner's dilemma inherent in their non-cooperative trade policies. In the unconstrained benchmark, a *shallow* agreement that focuses solely on tariff elimination is sufficient: once the technologically advanced country eliminates its tariff, the technologically backward economy voluntarily removes its export constraint. This conclusion changes when previous tariff liberalization has already reduced the scope for further tariff concessions. When tariffs are already sufficiently low, the technologically advanced economy has little incentive to offer further reductions if the technologically backward economy would unilaterally eliminate its export constraint even without additional tariff concessions. A second limitation of shallow agreements arises when we relax the simplifying assumption that the technologically backward economy can expand intermediate exports at no cost once the quantitative constraint is lifted. If export-licensing costs are sufficiently high, the backward economy no longer has an incentive to remove the export constraint in response to tariff cuts. In such cases, a trade agreement limited to tariff elimination offers no mechanism for inducing the technologically backward economy to liberalize its quantitative restriction, and therefore is not attractive to the technologically advanced country.

Achieving gains for both countries may instead require a *deep* trade agreement that disciplines quantitative export restrictions and thereby removes the export constraint in the technologically backward economy. Successful implementation of such an agreement leads to additional trade along the global value chain, a pattern that is well documented in recent empirical work (see Orefice and Rocha 2014; Johnson and

Noguera 2017; Osnago et al. 2017, 2019; Laget et al. 2020).[4] These insights, and the resulting need for a deep agreement to secure free trade, extend unchanged to the two-sector versions of the model, both with and without inter-sectoral labor mobility.

Our analysis, which focuses on global value chains, is closely related to the theoretical literature on vertical specialization and offshoring. Foundational contributions include Yi (2003), Kohler (2004), Grossman and Rossi-Hansberg (2008), and, more recently, Antràs and Chor (2013) and Costinot et al. (2013), who develop core models of sequential production in global value chains. Subsequent research has extended these frameworks in several directions. For example, Melitz and Redding (2014) study trade liberalization in intermediate goods and show that productivity gains from reorganizing sequential production can generate welfare improvements that are potentially unbounded. Zi (2020) examines how changes in trade costs for intermediates affect the distribution of rents between more and less developed countries along the value chain, while Costinot et al. (2012) analyze the distributional consequences for different types of labor throughout the production sequence.[5] Complementing this work, Kreickemeier and Qu (2020) explore the implications of final-goods trade liberalization and highlight the role of firm-level productivity differences in shaping global value chain organization. Additionally, Alfaro et al. (2019) develop a property-rights model that endogenizes firms' in-house production decisions, and Herkenhoff et al. (2024) study how corporate social responsibility evolves along the global value chain.

More closely related to our analysis, Antràs and Staiger (2012a), Antràs and Staiger (2012b), Antràs et al. (2024), and Blanchard et al. (2026) examine the role of value chain linkages in non-cooperative and cooperative trade policies, while Grossman et al. (2024) study how unanticipated tariff changes shape the design of global value chains (see Antràs and Chor 2022, for a review of this literature). Unlike these contributions, we focus on how the effectiveness of tariff-based and quantity-based trade policy instruments differs systematically along the global value chain. In particular, we show that while tariffs can be effective tools for downstream, technologically advanced economies, they may be ineffective for upstream, technologically backward economies, for which quantitative export constraints can emerge as a relevant policy instrument. In a recent paper, Lashkaripour and Lugovskyy (2023) characterize optimal trade policy in the presence of input-output linkages, allowing for both export and import tariffs (taxes or subsidies). While our primary aim is not to derive optimal trade policy, we show within a model of sequential production that tariffs can be ineffective for the technologically backward economy. This observation suggests that the set of available instruments may need to be expanded when considering optimal non-cooperative trade policy along the global value chain.[6]

[4] Larch and Yotov (2025) show in a structural gravity model that deep trade agreements may also foster foreign direct investment. In their framework, based on Anderson et al. (2020), foreign investment enters the foreign capital stock and therefore does not directly relate to trade along the value chain of final-goods production.

[5] This focus on distributional consequences is related to a broader literature on income distribution and productivity in open economies; see, for example, Hsu et al. (2023), who study how income inequality affects firm selection, aggregate productivity, and the gains from international trade.

[6] This distinguishes our analysis from Beshkar and Lashkaripour (2020), Caliendo and Parro (2022), and Caliendo et al. (2023), who examine optimal trade policy under a restricted set of instruments.

Because national policy interventions generate a terms-of-trade externality in our framework, our setting belongs to the class of models for which Bagwell and Staiger (2001) and Bagwell et al. (2016) show that cooperative tariff setting is sufficient to achieve a first-best trade equilibrium, provided that trade agreements also rule out the manipulation of other policies that distort trade. Consistent with this insight, our findings highlight that tariff elimination alone may be insufficient to make trade agreements attractive to both partners if it does not simultaneously prohibit quantitative export constraints, an instrument that is typically addressed only in deep trade agreements. In this respect, our analysis also relates to recent work by Grossman et al. (2021), who emphasize that even under free trade agreements countries may retain incentives to engage in regulatory protectionism, and that efficient agreements can be difficult to implement because they must discipline a broad set of non-tariff policies in a highly detailed manner.[7] In our setting, difficulties in achieving mutually beneficial free trade agreements arise instead from the presence of quantitative export constraints whose removal may be costly for one trading partner, combined with the absence of transfer mechanisms that would allow these costs to be efficiently shared.

Finally, our analysis relates to the "*latecomers*" problem emphasized by Bagwell and Staiger (2013). They argue that special and differential treatment allowed developing countries to benefit from MFN tariff reductions negotiated among developed countries without offering reciprocal concessions of their own, but also limited their influence over the pattern of developed-country liberalization. By the time developing countries were expected to participate more actively in reciprocal bargaining, developed-country tariffs in many sectors had already been reduced substantially, leaving little remaining bargaining power with which to bring developing countries to the table. This perspective also connects to recent work by Noël (2026) on the interaction between free trade agreements and global tariff liberalization. Studying a tariff-setting game with endogenous trade agreements, Noël (2026) shows that sectoral composition can determine whether free trade agreements facilitate or hinder multilateral tariff liberalization. Our paper identifies a complementary mechanism that arises in global value chains. We interpret an ex ante upper bound on the technologically advanced economy's tariff as the tariff policy space left by earlier liberalization, such as WTO tariff commitments. When this bound is sufficiently low, the technologically backward economy already removes its export quantity constraint unilaterally, and further tariff reductions are mainly redistributive. When the bound remains sufficiently high, however, there may still be an inefficient export constraint to negotiate over, so that a deep agreement linking tariff reductions to the removal of this constraint can be mutually beneficial. Thus, the extent of prior multilateral tariff liberalization shapes whether subsequent cooperation along global value chains remains feasible: when enough tariff bargaining space remains, a mutually beneficial deep agreement may still be possible;

[7] Related contributions also study when efficient trade agreements must extend beyond tariffs. Campolmi et al. (2025) analyze the efficiency of shallow versus deep agreements in a two-sector model with monopolistic competition, showing that efficiency requires balancing firm-entry effects and terms-of-trade considerations and therefore access to both trade and domestic policy instruments. Maggi and Ossa (2023) study cooperation over regulatory policies in a political-economy environment with lobbying.

when prior liberalization has already reduced tariffs too far, it can exhaust the tariff margin needed to make such cooperation mutually beneficial.[8]

The remainder of the paper is organized as follows. Section 2 presents empirical evidence that motivates our theoretical analysis. Sections 3 and 4 introduce the model and characterize the open economy equilibrium in the absence of policy distortions. Section 5 analyzes trade policy, highlighting the relative effectiveness of tariff- and quantity-based instruments. We show that import tariffs on intermediates and export tariffs on final goods benefit the technologically advanced economy, whereas export tariffs on intermediates and import tariffs on the final good are ineffective for the technologically backward economy, which may instead gain from imposing a quantitative export constraint on intermediates. Section 5 also examines the comparative-static effects of alternative instruments, studies the resulting non-cooperative policy game, and evaluates the scope for mutually beneficial trade agreements. Section 6 extends the model to a two-sector setting, while Section 7 considers the case in which lifting quantitative constraints and expanding intermediate exports is costly for the technologically backward economy. Section 8 concludes.

2 Motivating empirical evidence

We interpret the trade policy choices in the model outlined in the next section as the outcome of national-welfare-maximizing governments, and we present empirical evidence here to illustrate environments in which the incentives captured by the model are likely to be relevant. Specifically, we highlight two stylized facts: *(i)* the evolution of global value chain trade alongside the proliferation of shallow and deep trade agreements over time, and *(ii)* the differential association of these agreements with value-added versus global value chain trade. These stylized facts speak directly to the key mechanisms in our model and complement the broader observation, noted in the introduction, that quantitative export constraints have become increasingly important in input trade and that a large share of deep trade agreements contains provisions aimed at disciplining their use.[9]

[8] This perspective also complements the standard multi-sector logic of reciprocal tariff bargaining. In canonical settings, a mutually beneficial bargain over tariff policy often requires concessions across sectors or countries rather than reciprocity industry by industry (see, e.g., Bown and Sykes 2017). Sequential production creates a different possibility: because policies at different stages of the same value chain are complementary, both countries may have something to gain from liberalization within a single supply-chain relationship. Prior tariff liberalization can weaken this within-chain bargaining scope by making the remaining tariff negotiation primarily redistributive. If removing export restrictions is costly for the upstream country, deep provisions that directly discipline such restrictions can restore the scope for mutually beneficial cooperation.

[9] Deep agreements include provisions covering a wide range of policy areas such as competition policy, investment, corruption, and intellectual property rights (see Hofmann et al. 2019; Laget et al. 2020). In a systematic evaluation of the content of deep trade agreements, Wu (2020) highlights the widespread use of provisions disciplining quantitative export constraints. While the purpose of this section is to shed light on the role of deep trade agreements in the context of global value chain trade, we do not aim to contribute to the broader debate on whether shallow or deep trade agreements have stronger effects on aggregate trade flows. That question has been studied elsewhere, with mixed findings (see, e.g., Mattoo et al. 2022; Besedeš et al. 2025). Our focus is instead on the role of export quantity constraints and on how provisions disciplining their use are associated with value-added and global value chain trade.

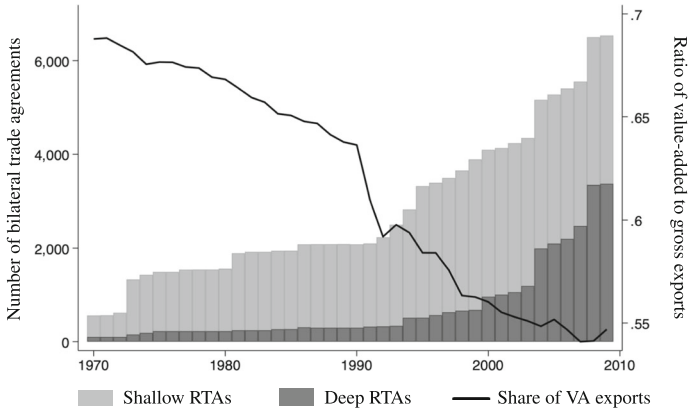


Fig. 1 Evolution of value-added trade and the number of trade agreements. *Notes:* Data sources are the WTO database on deep trade agreements, the CEPII gravity database, and the replication data of Johnson and Noguera (2017). The reported export data covers production sectors only

We begin by documenting long-run trends in value-added and global value chain trade. Using the data of Johnson and Noguera (2017), we find that between 1970 and 2010 the share of value-added trade fell by roughly 15 percentage points, indicating a substantial rise in global value chain trade during this period. Figure 1 also reports the evolution of trade agreements over the same period, drawing on the WTO database for deep trade agreements and the CEPII gravity database (see Hofmann et al. 2017; Conte et al. 2021, for further details). We distinguish between shallow and deep regional trade agreements (RTAs), with the latter covering a broader set of provisions beyond tariff reductions. The number of bilateral RTAs has increased markedly, with a particularly pronounced rise in deep agreements since the mid-1990s. We interpret this pattern as suggestive of a co-movement between the expansion of global value chain trade and the proliferation of free trade agreements, with the former preceding (and potentially motivating) the latter.

To complement these descriptive patterns, we examine how the presence of different types of trade agreements correlates with bilateral value-added and global value chain trade flows. We use data from the EORA multi-region input-output tables, covering industry-level bilateral exports for a balanced sample of 184 countries from 1998 to 2015. The input-output structure allows us to distinguish between value-added (VA) and global value chain (GVC) trade, with the latter capturing flows crossing more than one border (Belotti et al. 2021). We relate these trade measures to the presence of shallow and deep RTAs using a standard structural gravity specification estimated by Poisson pseudo maximum likelihood, controlling for country-pair, exporter-time, and importer-time fixed effects (see Larch and Yotov 2024).

Panel A of Figure 2 reports the estimated correlations. Both shallow and deep agreements are positively associated with value-added exports, whereas only deep agreements are associated with higher GVC exports (significant at the 10 percent level). Panel B distinguishes deep agreements with and without provisions disciplining quantitative export constraints. As in Panel A, all agreement types correlate positively

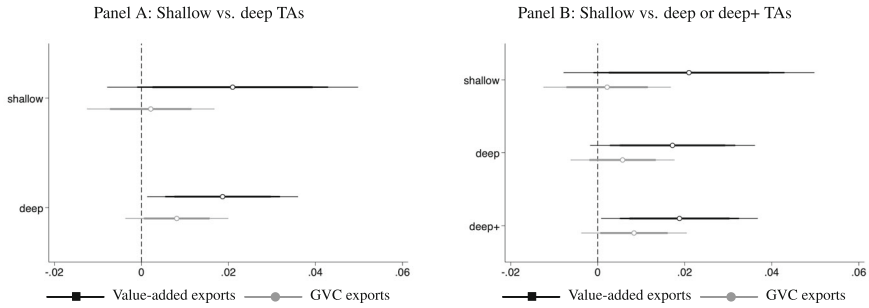


Fig. 2 The impact of shallow and deep trade agreements on bilateral trade flows. *Notes:* Data sources are the EORA multi-region input-output tables, the WTO database on deep trade agreements, and the CEPII gravity database for the years 1998-2015. The dependent variable is bilateral exports in millions of current US dollars. Trade agreements are classified as *shallow*, *deep*, or *deep+*, with the latter explicitly disciplining the use of quantitative export constraints. The regressions include country-pair fixed effects as well as importer-time and exporter-time fixed effects. Standard errors are clustered at the country-pair level. Thick, medium, and thin lines denote the 90, 95, and 99 percent confidence intervals, respectively

with value-added trade, but only deep agreements containing such provisions (denoted deep+) are positively associated with GVC trade (again at the 10 percent level).

We interpret these correlations as suggestive that quantitative export constraints dampen the fragmentation of production across borders, and that deep agreements disciplining such constraints serve as an important catalyst for expanding trade along the global value chain.

3 Basic model structure

We conduct our formal analysis using the production model of Costinot et al. (2013), where a continuum of stages has to be completed to produce a single consumption good. Assuming that these stages must be executed consecutively in a predetermined order, the model represents a value chain, which we assume to have a unit length. In our analysis, each production stage is represented by an index s from the unit interval.

We integrate this production model into a world consisting of two countries, $c = 1, 2$, each endowed with an equal mass of labor L but differing in their production technologies. Labor is immobile between countries, and technology determines the Poisson rate, $\lambda_c \in (0, 1)$, at which mistakes occur at each production stage. Whenever a mistake happens, the processed intermediate is destroyed, resulting in a wasteful loss of labor input from all previous stages. We assume that mistakes occur at a higher rate in country 1 than in country 2, i.e. $\lambda_1 > \lambda_2$. Consequently, country 1 is the technologically backward economy (the "South"), while country 2 is the technologically advanced economy (the "North").

Considering a Leontief technology that combines $q(s)$ units of the intermediate produced at stage s with $q(s)ds$ units of labor to produce $q(s + ds)$ units at the consecutive stage $s + ds$, we can express the production technology at stage s as $q(s + ds) = (1 - \lambda_c ds)q(s)$. For infinitesimal ds, the production function of country

c at stage s can be written as a differential equation of the following form:

$$q'(s) = -\lambda_c q(s). \tag{1}$$

Markets at all stages are perfectly competitive. The world price of the intermediate good at stage s is given by $p(s)$. The initial input at stage $s = 0$ is available in perfectly elastic supply at a price of $p(0) = 0$. The final good at stage $s = 1$ is chosen as our numéraire, with its (producer) price normalized to one. The common wage paid in country c is denoted by w_c.[10]

Using our technology assumption, the costs for country c to produce output $q(s+ds)$ are $p(s)q(s) + w_c q(s)ds$. Substituting $q(s + ds) = (1 - \lambda_c ds)q(s)$ gives the unit cost function for stage s as: $[p(s) + w_c ds]/(1 - \lambda_c ds)$. Under perfect competition, profit-maximization establishes:

$$\begin{aligned} p(s + ds) &\leq \frac{p(s) + w_c ds}{1 - \lambda_c ds}, \\ p(s + ds) &= \frac{p(s) + w_c ds}{1 - \lambda_c ds} \quad \text{if} \quad Q_c(s') > 0 \quad \text{for all} \quad s' \in (s, s + ds], \end{aligned} \tag{2}$$

where $Q_c(s)$ is the economy-wide output of stage s in country c. Labor market clearing for country c implies:

$$\int_0^1 Q_c(s)ds = L. \tag{3}$$

This completes the description of the model structure.

4 The open economy without policy distortions

In this section, we investigate a baseline version of our model in the absence of policy, reducing our setting to a two-country variant of the model proposed by Costinot et al. (2013). The open economy equilibrium is characterized by the following lemma.

Lemma 1 *In the open economy equilibrium, there exists a unique partitioning of the unit interval of production stages into two subintervals of length S_1 and $1 - S_1$, respectively, such that $Q_1(s) > 0$ if and only if $s \in (0, S_1]$ and $Q_2(s) > 0$ if and only if $s \in (S_1, 1]$. We have:*

$$S_1 = -\frac{1}{\lambda_1} \ln\left(1 - \frac{\lambda_1 L}{Q_0}\right), \qquad 1 = S_1 - \frac{1}{\lambda_2} \ln\left(1 - \frac{\lambda_2 L}{Q_1}\right), \tag{4}$$

[10] To apply the solution concept of Costinot et al. (2013), we impose the additional formal condition that each firm produces a measure $\Delta > 0$ of consecutive stages. Specifically, if a firm produces stage $s' = s+ds$, it produces all stages $s' \in (\underline{s}, \underline{s} + \Delta]$. This implies that each unit of the consumption good is produced by a finite number of firms.

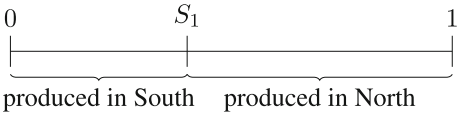


Fig. 3 The distribution of production stages across countries

and output levels:

$$Q_1 = Q_0 - \lambda_1 L, \qquad Q_2 = Q_1 - \lambda_2 L, \tag{5}$$

with $Q_0 \equiv Q_1(0)$, $Q_1 \equiv Q_1(S_1) = Q_2(S_1)$, *and* $Q_2 \equiv Q_2(1)$.

Proof The first part of the lemma on sequential production follows from the proof of Proposition 1 in Costinot et al. (2013). To derive Eqs. (4) and (5), we solve differential equation (1) at the country level and compute the general solutions $Q_1(s) = \exp[-\lambda_1 s]Q_0$ and $Q_2(s) = \exp[-\lambda_2(s - S_1)]Q_1$. Substituting these two solutions into the labor market clearing conditions of countries 1 and 2 establishes Eq. (4). Eq. (5) then follows from substituting Eq. (4) into $Q_1 = \exp[-\lambda_1 S_1]Q_0$ and $Q_2 = \exp[-\lambda_2(1 - S_1)]Q_1$. □

Figure 3 provides a graphical representation of the global value chain in our model. The key finding of Lemma 1 is that the South executes the initial stages of the global value chain, while the North executes the later stages. This is a direct consequence of our assumption that country 1 is the technologically backward economy, and the outcome is intuitive because a mistake at stage $s > 0$ destroys the production from all previously executed stages. Hence, the loss from mistakes and the (expected) costs of production are minimized by the sequential production structure depicted in Figure 3.

The two equations in (4) can be combined to form the global labor market clearing condition. Introducing an auxiliary variable to measure the level of initial input relative to effective Southern labor supply, $x \equiv Q_0/(\lambda_1 L)$, and denoting the relative productivity gap between the South and the North by $\Lambda \equiv \lambda_2/\lambda_1 < 1$, we can express the global labor market clearing condition as follows:

$$F(x; \Lambda) \equiv -\ln\left(\frac{x-1}{x}\right) - \frac{1}{\Lambda}\ln\left(\frac{x-1-\Lambda}{x-1}\right) - \lambda_1 = 0, \tag{6}$$

where $Q_1/(\lambda_1 L) = x - 1$ has been substituted from Eq. (5). Eq. (6) implicitly determines the equilibrium level of x in the open economy. Noting that $\lim_{x\to 1+\Lambda} F(x, \Lambda) = \infty$, $\lim_{x\to\infty} F(x, \Lambda) = -\lambda_1$, and $F'_x(x; \Lambda) < 0$ it follows that the solution for x is unique and satisfies $x > 1 + \Lambda$. We can further compute $F'_\Lambda(x; \Lambda) > 0$. Applying the implicit function theorem, we therefore obtain $dx/d\Lambda > 0$: For a given λ_1, a lower productivity disadvantage of the South, captured by a higher value of Λ, is associated with a larger mistake rate in the North. This requires higher initial labor input and consequently a decline in the fraction of pro-

duction stages executed in the South to allow for full employment of workers along the global value chain.[11]

With the characterization of the open economy equilibrium at hand, we continue our analysis by determining the welfare of North and South in the open economy. Since the single consumption good serves as our numéraire, the wage rate paid in country c is equal to the indirect utility of consumers in country c: $v_c = w_c$. To determine welfare in the open economy equilibrium, we compute the equilibrium levels of w_1 and w_2, which are jointly determined with the price of traded intermediates $p_1 \equiv p(S_1)$ by the following three-equation system:

$$\begin{aligned} \lambda_1 p_1 &= \{\exp[\lambda_1 S_1] - 1\}\, w_1, \\ \lambda_1 &= \exp[\lambda_1 \Lambda(1 - S_1)]\lambda_1 p_1 + \{\exp[\lambda_1 \Lambda(1 - S_1)] - 1\}\, \frac{w_2}{\Lambda}, \\ w_2 &= w_1 + \lambda_1 p_1 (1 - \Lambda). \end{aligned} \tag{7}$$

The derivation of (7) is tedious and thus deferred to Appendix A.1. However, the three equations have a straightforward economic interpretation. Using Lemma 1, we find that the first line corresponds to the binding budget constraint of country 1, while the second line corresponds to the binding budget constraint of country 2. The third line captures a no-arbitrage condition under the profit-maximizing choice of executing stages $s \in (0, S_1]$ in country 1 and stages $s \in (S_1, 1]$ in country 2. The equation system (7) can be solved for:

$$\begin{aligned} w_1 &= \lambda_1 \frac{(x-1)(x-1-\Lambda)}{2x-1-\Lambda}, \quad w_2 = \lambda_1 \frac{(x-\Lambda)(x-1-\Lambda)}{2x-1-\Lambda}, \\ p_1 &= \frac{x-1-\Lambda}{2x-1-\Lambda}. \end{aligned} \tag{8}$$

It is easily confirmed from Eq. (8) that $w_1 = w_2$ if $\Lambda = 1$, whereas $w_1 < w_2$ whenever $\Lambda < 1$. This shows that the technologically advanced North pays higher wages and is therefore better off than the technologically backward South. In the next section, we analyze how trade policy affects the open economy equilibrium.

5 The open economy with policy distortions

In this section, we analyze the role of trade policy in shaping the structure of sequential production and the distribution of economic rents between the South and the North. We begin by examining the effectiveness of import and export tariffs along the global value chain in Subsect. 5.1. In Subsect. 5.2, we extend the set of trade policy instruments,

[11] With x increasing in Λ, $F(x, 1) = 0$ determines an upper bound for x, denoted by $\overline{x}$, whereas $F(x, 0) = 0$ gives a lower bound of x, denoted by $\underline{x}$. Evaluated at $\Lambda = 1$, Eq. (6) establishes $\overline{x} = 2\exp[\lambda_1]/\{\exp[\lambda_1]-1\}$. Moreover, we compute $\lim_{\Lambda\to 0} F(x, \Lambda) = -\ln[(x-1)/x] + (x-1)^{-1} - \lambda_1 \equiv \Phi(x)$, with $\Phi'(x) < 0$ and $\lim_{x\to\infty} \Phi(x) = -\lambda_1$. Acknowledging $\Phi(2.79) > 1 - \lambda_1 > 0$, it follows that $\underline{x} > 2.79$ holds for all possible realizations of λ_1.

adding a quantitative constraint that reduces exports of intermediates from the South to the North.

5.1 The effectiveness of import and export tariffs along the value chain

We first consider arbitrary levels of ad-valorem export and import tariffs for the two economies in the one-sector model outlined in Sections 3 and 4. We denote the tariff imposed by country $c \in \{1, 2\}$ on intermediates by $\tau_c > 0$ and its tariff on the final good by $\zeta_c > 0$. A tariff rate lower (higher) than one refers to a subsidy (tax), whereas a rate equal to one reflects zero tariffs. We assume that tariff revenues are rebated lump sum to households and that any public expenditures are financed by lump-sum taxation. Under these assumptions, the structure of sequential production and the global labor market clearing condition in Eq. (6) remain unchanged when $\zeta_c, \tau_c \neq 1$. The reason is that the Leontief technology eliminates substitution effects: relative price changes induced by tariffs do not alter the composition of inputs across production stages. Consequently, the equilibrium value of the initial input per effective unit of Southern labor, $x = Q_0/(\lambda_1 L)$, is invariant to all tariff instruments. Appendix A.2 derives, analogous to Section 4, a system of three equations determining wages and prices in the open economy, which solves to

$$w_1 = \frac{\lambda_1}{\tau_1 \tau_2} \frac{(x-1)(x-1-\Lambda)}{2x-1-\Lambda}, \quad w_2 = \lambda_1 \frac{(x-\Lambda)[x-1-\Lambda]}{2x-1-\Lambda},$$
$$p_1 = \frac{1}{\tau_1 \tau_2} \frac{x-1-\Lambda}{2x-1-\Lambda}. \qquad (8')$$

where p_1 is the free-on-board price of Q_1, excluding export tariffs of country 1, τ_1, and import tariffs of country 2, τ_2.

To determine welfare in the open economy for arbitrary tariff levels, we introduce auxiliary variables C_Q^1 and C_Q^2 denoting final-good consumption in the two countries. With the producer price of the (numéraire) final good normalized to one, consumption expenditures in the South equal $\zeta_1 \zeta_2 C_Q^1$, where ζ_1 and ζ_2 are, respectively, the South's import tariff and the North's export tariff on final goods. A balanced budget in country 1 requires $\zeta_1 \zeta_2 C_Q^1 = w_1 L + (\tau_1 - 1) p_1 Q_1 + (\zeta_1 - 1)\zeta_2 C_Q^1$. Rearranging terms, per-capita consumption in the South can be expressed as:

$$\frac{C_Q^1}{L} = \frac{w_1}{\zeta_2} + \frac{\tau_1 - 1}{\zeta_2} \lambda_1 p_1 \frac{Q_1}{\lambda_1 L} = \frac{\lambda_1}{\tau_2 \zeta_2} \frac{(x-1)(x-1-\Lambda)}{2x-1-\Lambda}, \qquad (9)$$

where the second equality uses $Q_1/(\lambda_1 L) = x - 1$ and p_1 from Eq. (8′). Southern welfare is given by $v_1 = C_Q^1/L$, which is invariant to changes in the South's export tariff τ_1 and import tariff ζ_1 because, as noted above, the equilibrium value of x does not depend on the tariff instruments.

Next, we turn to welfare in the North, where consumers face a budget constraint of $C_Q^2 = w_2 L + (\tau_2 - 1)\tau_1 p_1 Q_1 + (\zeta_2 - 1) C_Q^1$. Rearranging terms, per-capita welfare

in the North can be expressed as:

$$\begin{aligned}\frac{C_Q^2}{L} &= w_2 + (\tau_2 - 1)\lambda_1\tau_1 p_1 \frac{Q_1}{\lambda_1 L} + (\zeta_2 - 1)\frac{C_Q^1}{L} \\ &= \lambda_1 \frac{(x-1)(x-1-\Lambda)}{2x-1-\Lambda}\left(\frac{x-\Lambda}{x-1} + \frac{\zeta_2\tau_2 - 1}{\zeta_2\tau_2}\right),\end{aligned} \tag{10}$$

where the second equality uses $Q_1/(\lambda_1 L) = x - 1$ and $p_1, C_Q^1/L$ from Eqs. (8′) and (9), respectively. Northern welfare is given by $v_2 = C_Q^2/L$, which shows that τ_2 and ζ_2 are complementary policy instruments for the North. We can therefore summarize the main insights from our analysis in this subsection with the following lemma:

Lemma 2 *In a one-sector model of sequential production with a Leontief technology, neither export tariffs on intermediates nor import tariffs on the final good are effective trade policy instruments for the South. In contrast, the more advanced North, which executes the final stages of production, finds both import tariffs on intermediates and export tariffs on the final good to be equally effective trade policy instruments.*

Proof Analysis in the text and formal derivation details in Appendix A.2. □

The analysis in this section reveals that in the case of sequential production, the technologically backward economy cannot use tariff instruments to appropriate a larger share of the economic rents produced along the global value chain. This is different for the technologically advanced economy, which can use tariffs for this purpose. Due to this asymmetry in the effectiveness of conventional tariff-based trade policy instruments, the South must rely on alternative, quantity-based measures. For instance, the South can reduce the amount of intermediates shipped to the North for further processing. However, this is attractive for the South only if the non-exported intermediates are not pure waste but generate (at least some) positive welfare for local consumers. We explore this possibility in the next subsection.

5.2 Tariff-based and quantity-based trade policy instruments

We now expand the set of policy instruments for the South by allowing it to use a quantity instrument that limits the export volume of intermediates by setting an upper bound to the fraction of workers employed in the production of intermediates to $\delta \in [0, 1]$. With a fraction $\delta \leq 1$ of Southern labor supply used in the production of exported intermediates, we find — under a slight abuse of notation — x to be given by $Q_0/(\delta\lambda_1 L)$. Following the analysis in Section 4, we can then characterize the profit-maximizing structure of global value chain production in the open economy by the unchanged upper threshold for the global value chain stages produced in country 1, $\lambda_1 S_1 = -\ln[(x-1)/x]$ and the modified labor market clearing condition for workers employed in global value chain production:

$$\tilde{F}(x;\Lambda,\delta) \equiv -\ln\left(\frac{x-1}{x}\right) - \frac{1}{\Lambda}\ln\left(\frac{\delta(x-1)-\Lambda}{\delta(x-1)}\right) - \lambda_1 = 0. \tag{6′}$$

Eq. (6′) pins down the initial input per unit of effective Southern labor used in global value chain production, x, as a function of exogenous relative labor productivity Λ and the export constraint δ. The derivation properties of $\tilde{F}(x;\Lambda,\delta)$ with respect to x and Λ are inherited from $F(x;\Lambda)$, implying that $dx/d\Lambda > 0$ continues to hold. Additionally, noting $\partial\tilde{F}(\cdot)/\partial\delta < 0$, we can conclude that $dx/d\delta < 0$. To acknowledge the impact of changes in policy parameter δ on the initial output of effective labor input, we consider $x = x(\delta)$, with $x'(\delta) < 0$, from now on. While increasing the initial input per effective Southern labor use x, a stronger binding export constraint lowers the quantity of intermediates shipped from the South to the North. This is captured by a positive impact of δ on the product $\delta(x-1)$, an effect that is formally shown in Appendix A.3. A lower δ therefore reduces Northern labor demand. Thus, full employment in the North and the employment of a fraction δ of Southern labor require adjustments of the global value chain, which are captured by a decline in the fraction of production stages executed in the South, S_1.

The residual fraction $1-\delta$ of Southern labor supply is employed in the local production of the consumption good in an outside sector using an inferior technology, which is captured by the purely national value chain used under autarky. Therefore, market clearing in the outside sector is given by $\lambda_1 = -\ln\left[(x^a-1)/x^a\right]$, where $x^a \equiv Q_0^a/[(1-\delta)\lambda_1 L]$ and superscript a is added to distinguish the autarkic production of the consumption good from fragmented production along the global value chain. The market outcome is efficient at $\delta = 1$, so a binding quantity constraint that reduces employment in global value chain production is distortionary from a social planner's point of view. Setting $\delta < 1$ reduces global production of the homogeneous consumption good, which equals the joint output of global value chain production Q_2 and autarkic production $Q^a \equiv Q_0^a - (1-\delta)\lambda_1 L$.

To determine the welfare effects of changing tariff- and quantity-based policy instruments, we derive an equation system similar to (7) and solve for wages, w_1 and w_2, and the (free-on-board) export price, p_1 along the global value chain. As formally shown in Appendix A.2, we compute:

$$w_1 = \frac{\lambda_1}{\tau_1\tau_2}\frac{(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda}, \quad w_2 = \lambda_1\frac{(x-\Lambda)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda},$$
$$p_1 = \frac{1}{\tau_1\tau_2}\frac{\delta(x-1)-\Lambda}{\delta(x-1)+x-\Lambda}. \tag{8″}$$

Similar to our analysis from the previous section, both a higher Northern import tariff and a higher Southern export tariff lower Southern wages w_1 while leaving the Northern wages w_2 unaffected. With the amount of initial input per effective Southern labor use $x = x(\delta)$ — and thus the production structure along the global value chain — unchanged, a higher import tariff on intermediates is purely redistributive. This is reflected by a decline in the Southern export price p_1, which improves the North's terms of trade. In contrast, a stronger Southern export constraint, i.e. a lower level of δ, is distortionary and reduces the overall production output of the consumption good. As formally shown in Appendix A.3, a lower δ causes wages in global value chain production to decrease in the North and to increase in the South and results in a decline in the export price of intermediates.

Regarding the $1-\delta$ units of labor used in the autarkic production of the consumption good, we assume that they produce an output that is sold under the same local price as good Q_2. In the presence of final goods tariffs the price of good Q^a is therefore given by $\zeta_1\zeta_2$, while wage paid to workers employed in autarkic production is given by $w_1^a = \zeta_1\zeta_2\lambda_1/\{\exp[\lambda_1]-1\}$. Accordingly, the real wage of workers employed in the outside sector corresponds to $\lambda_1/\{\exp[\lambda_1]-1\}$ and it is independent of both price- and quantity-based trade policy instruments. Denoting economy-wide consumption of the autarkic good by $C_Q^a \equiv Q^a$, the budget constraint of the South is given by $\zeta_1\zeta_2 C_Q^1 + \zeta_1\zeta_2 C_Q^a = \delta w_1 L + (\tau_1 - 1)p_1 Q_1 + (\zeta_1 - 1)\zeta_2 C_Q^1 + (1-\delta)w_1^a L$ and it solves for

$$\frac{C_Q^1}{L} = \frac{\delta w_1}{\zeta_2} + \frac{\tau_1 - 1}{\zeta_2}\lambda_1 p_1 \frac{Q_1}{\lambda_1 L} = \frac{\lambda_1}{\zeta_2\tau_2}\frac{\delta(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda},$$
$$\frac{C_Q^a}{L} = \frac{(1-\delta)\lambda_1}{\exp[\lambda_1]-1}. \tag{9'}$$

Welfare in the South is then given by $v_1 = C_Q^1/L + C_Q^a/L$. Similarly, the budget constraint in the North is given by $C_Q^2 = w_2 L + (\tau_2 - 1)\tau_1 p_1 Q_1 + (\zeta_2 - 1)C_Q^1$, which establishes (following the derivation steps from the previous subsection):

$$\frac{C_Q^2}{L} = \lambda_1 \frac{\delta(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda}\left(\frac{x-\Lambda}{\delta(x-1)} + \frac{\zeta_2\tau_2 - 1}{\zeta_2\tau_2}\right), \tag{10'}$$

with Northern welfare given by $v_2 = C_Q^2/L$.

Contrasting Eqs. (9′) and (10′) with their respective counterparts in Subsect. 5.1, we conclude that the findings from Lemma 2 extend to the case where the South imposes an export quantity constraint on intermediates. In particular, Southern tariffs are ineffective policy instruments, so we abstract from them in the subsequent analysis by setting $\tau_1 = \zeta_1 = 1$. Moreover, Northern import tariffs on intermediates and Northern export tariffs on final goods are complements and exert qualitatively similar effects on welfare in the North and the South. We therefore set $\tau \equiv \zeta_2\tau_2$ and $\zeta_2 = 1$, so that we can refer to τ as a Northern import tariff on intermediates and keep our discussion simple.

Focusing on the two policy instruments δ and τ, we can express Southern welfare in the following form:

$$v_1 = \frac{\lambda_1}{\tau}\frac{\delta(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} + \frac{(1-\delta)\lambda_1}{\exp[\lambda_1]-1} \equiv \tilde{v}_1(\tau,\delta). \tag{11}$$

Similarly, Northern welfare equals

$$v_2 = \lambda_1\frac{(x-\Lambda)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} + \lambda_1\frac{\tau-1}{\tau}\frac{\delta(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} \equiv \tilde{v}_2(\tau,\delta), \tag{12}$$

where $x = x(\delta)$ follows from Eq. (6′).[12] By setting δ, the Southern government imposes an upper bound on the share of workers employed in global value chain production. The welfare functions in Eqs. (11) and (12) have been derived under the implicit assumption that this upper bound is binding. However, the quantity constraint restricts labor mobility only in one direction, namely from autarkic production to global value chain production, but not vice versa. For δ to be decisive for the size of the export sector, it must therefore be the case that wages in global value chain production are at least as high as in autarkic production, i.e., $w_1 \geq w_1^a$. This requires the Northern tariff to be sufficiently low, i.e.

$$\tau \leq \{\exp[\lambda_1] - 1\} \frac{(x-1)[\delta(x-1) - \Lambda]}{\delta(x-1) + x - \Lambda} \equiv \overline{\tau}(\delta), \tag{13}$$

according to the solution for w_1 in Eq. (8″) and the solution for w_1^a in the text. From our previous insight that $dw_1/d\tau < 0$ and $dw_1/d\delta < 0$, it follows that $\overline{\tau}'(\delta) < 0$, with $\overline{\tau}(\delta)$ reaching its maximum level at $\delta = 0$:

$$\lim_{\delta \to 0} \overline{\tau}(\delta) = \frac{\lambda_2/(\exp[\lambda_2] - 1)}{\lambda_1/(\exp[\lambda_1] - 1)} \equiv \overline{\tau} > 1.$$

In Appendix A.3, we show that $\partial \tilde{v}_1(\overline{\tau}, 0)/\partial\delta = 0$. The following lemma summarizes how unilateral changes in the two policy instruments affect welfare of the two trading partners in the open economy.

Lemma 3 *For any given $\delta > 0$, an increase in the Northern import tariff up to the threshold $\overline{\tau}(\delta)$ increases welfare in the North and reduces welfare in the South. For any $\tau < \overline{\tau}(\delta)$, a stricter Southern export constraint lowers welfare in the North, whereas the welfare effects in the South depend on the concrete realisation of the Northern import tariff and are not a priori clear. There exists a critical tariff $\underline{\tau}$, which is higher than 1, lower than $\overline{\tau}$, and implicitly determined by $\partial \tilde{v}_1(\underline{\tau}, 1)/\partial\delta = 0$, such that lowering δ below its maximum level of one decreases Southern welfare monotonically for all $\tau \leq \underline{\tau}$. For $\tau \geq \overline{\tau}$, lowering δ below its maximum level of one increases Southern welfare monotonically. Finally, Southern welfare has a unique interior maximum at some $\delta \in (0, 1)$ if $\tau \in (\underline{\tau}, \overline{\tau})$.*

Proof See Appendix A.3. □

Lemma 3 characterizes the best responses of the two countries in a non-cooperative policy game. The North prefers the highest feasible tariff rate, constrained by the incentive-compatibility condition that wages in global value chain production must not fall below autarky wages for Southern workers. Provided that $\delta > 0$, this makes $\tau = \overline{\tau}(\delta)$ the preferred choice for the North. If $\delta = 0$, the North is indifferent between all levels of τ, because there are no tariff revenues for the North if the South ceases to export intermediates.

[12] Making use of Eq. (8″), we can alternatively express welfare in the two countries in a more concise way as $\tilde{v}_1(\tau, \delta) = \delta w_1 + (1 - \delta) w_1^a$ and $\tilde{v}_2(\tau, \delta) = w_2 + (\tau - 1)\delta w_1$.

The South experiences two counteracting effects of lowering its export constraint. On the one hand, increasing δ shifts more workers towards global value chain production, provided that the initial level of $\delta < 1$ generates a binding mobility constraint for workers due to $w_1 > w_1^a$. On the other hand, increasing δ reduces the wage premium in global value chain production leading to an infra-marginal welfare loss of workers initially employed in the Southern export sector. Balancing these two counteracting effects leads to a welfare-maximizing level of $\delta \in (0, 1)$ if the North sets an intermediate level of import tariffs $\tau \in (\underline{\tau}, \overline{\tau})$. If Northern tariffs are moderate, i.e. if $\tau \leq \underline{\tau}$, the South eliminates the export constraint entirely and sets $\delta = 1$. This is attractive despite a positive level of the Northern import tariff, because the South has to bear part of the efficiency loss generated by the export quantity constraint itself. In contrast, if Northern tariffs exceed the critical threshold of $\overline{\tau}$, the South bans exports entirely by setting $\delta = 0$, because even at the minimum level of δ, wages in the export sector are lower than in autarkic production in this case.

Noting finally that the existence of infra-marginal Southern gains from lowering δ is sufficient for $\overline{\tau}(1) > \underline{\tau}$, it follows that for any $\delta > 0$ the best-response of the South to the preferred tariff choice of the North, $\tau = \overline{\tau}(\delta)$, is to reduce δ in order to realize a wage increase for the workers employed in global value chain production. This induces the North to further increase its import tariff in response to the stronger Southern export constraint. The mutual incentive to increase the own share of globalization rents leads to a Nash equilibrium in pure strategies characterized by the following proposition.

Proposition 1 *The non-cooperative policy equilibrium is given by* $\tau \geq \overline{\tau}$ *and* $\delta = 0$ *and establishes*

$$\lim_{\tau\to\infty} \tilde{v}_1(\tau, 0) = \frac{\lambda_1}{\exp[\lambda_1] - 1} \equiv v_1^{nc}, \quad \lim_{\tau\to\infty} \tilde{v}_2(\tau, 0) = \frac{\lambda_2}{\exp[\lambda_2] - 1} \equiv v_2^{nc} \tag{14}$$

Proof The proposition follows from Eqs. (11), (12), and Lemma 3. □

Non-cooperative policy setting results in an outcome where $\lim_{\delta\to 0} S_1 = 0$, implying that both countries use the entire labor supply in their autarkic production technologies. Consequently, non-cooperative policy setting puts both countries into their respective closed economy equilibrium. This is intuitive because, in response to the North setting an excessive import tariff, the South stops exporting intermediates, resulting in $v_1^{nc} = v_1^a$ and $v_2^{nc} = v_2^a$, with the superscript a referring to autarky. Since non-cooperative policy setting eliminates global value chain trade in our model, it is clear that cooperation of the two governments can be welfare-improving. To illustrate this, we contrast the non-cooperative welfare levels in Proposition 1 with the welfare levels that could be achieved if the two countries coordinated on free trade without export constraints, i.e. on $\delta = \tau = 1$. The results from this comparison are summarized in Proposition 2.

Proposition 2 *Policy coordination on* $\tau = \delta = 1$ *increases welfare in both countries relative to the non-cooperative case in Proposition 1.*

Proof See the Appendix. □

The preceding results establish that eliminating both tariff- and quantity-based impediments to trade yields an efficient open-economy equilibrium in our model. Since there are no market failures in this environment, removing all policy distortions maximizes global welfare (see Antràs and Chor 2022, for discussion). This efficiency statement, however, does not by itself imply that moving from the non-cooperative equilibrium to free trade is mutually beneficial. As shown by Johnson (1953), even in an environment with perfect competition and no domestic market failures, a country may prefer a non-cooperative tariff equilibrium to free trade if it can shift part of the tariff burden onto its trading partner. Subsequent work has shown that such gains from tariff retaliation depend importantly on country asymmetries and relative size (e.g., Kennan and Riezman 1988; Syropoulos 2002).

Against this background, Proposition 2 provides a strong benchmark result for our framework: despite the asymmetry between the two countries, coordination on $\tau = \delta = 1$ increases welfare in both countries relative to the non-cooperative equilibrium. This is not an immediate implication of global efficiency. In our value-chain environment, however, non-cooperation leads to autarky, as shown in Proposition 1. This stark outcome reflects the complementarity between Northern tariff policy and the Southern export constraint in sequential production. As Proposition 2 shows, the gains from restoring trade along the value chain are sufficiently large to make the efficient open-economy allocation mutually beneficial.

Lemma 3 explains why this outcome can be implemented through a shallow agreement in the unconstrained baseline case. Since $\underline{\tau} > 1$, tariff elimination by the North implies $\tau = 1 < \underline{\tau}$. The tariff is therefore in the range in which lowering δ below one reduces Southern welfare, so that the South's best response is to set $\delta = 1$ and eliminate the export quantity constraint. A shallow trade agreement that focuses solely on tariff elimination thus induces the South to remove the quantity constraint voluntarily if this removal is costless. Hence, in the unconstrained baseline case, a shallow agreement is sufficient both to restore global efficiency and to make both countries better off relative to non-cooperation.

However, this benchmark relies on the assumption that the North is unconstrained in its tariff choice. In practice, tariff policy is typically bounded by institutional or legal commitments, such as WTO obligations. To capture this feature, we now consider the case in which the North's import tariff is subject to an upper bound, denoted by $\tau_{\max}$. The following proposition characterizes the implications of such a constraint.

Proposition 3 *If Northern import tariffs are constrained by an upper bound $\tau_{\max}$, a trade agreement that imposes $\tau = \delta = 1$ is beneficial for both countries only if $\tau_{\max}$ is sufficiently high. If $\tau_{\max}$ is too low, a trade agreement on $\tau = \delta = 1$ is not beneficial for the North.*

Proof Analysis in the text. □

Proposition 3 follows directly from the best-response policies characterized in Lemma 3. If $\tau_{\max} \geq \overline{\tau}$, the non-cooperative equilibrium features $\tau \geq \overline{\tau}$ and $\delta = 0$, and cooperation on $\tau = \delta = 1$ is mutually beneficial, as in Proposition 2. By contrast, if $\tau_{\max} \leq \underline{\tau}$, the non-cooperative equilibrium is characterized by $\tau = \tau_{\max}$ and $\delta = 1$. In this case, the South already eliminates its export constraint unilaterally. Further tariff

reductions therefore benefit the South but harm the North, because tariffs are purely redistributive in the baseline model. By continuity, there exists an intermediate range of tariff bounds in which mutually beneficial cooperation is possible only if $\tau_{\max}$ is sufficiently high.

The low-tariff case highlights an important limitation of the baseline environment. When $\tau_{\max} \leq \underline{\tau}$, the tariff liberalization already embodied in the ex ante bound leaves no scope for a mutually beneficial trade agreement, whether shallow or deep. A shallow agreement, which operates only through further tariff reductions, merely redistributes surplus from the North to the South. A deep agreement cannot generate additional gains either: although it could, in principle, also discipline the South's export quantity constraint, this additional margin is irrelevant in the low-tariff region because the South has already removed the constraint in the non-cooperative equilibrium. Hence, absent cross-border transfers, any agreement that entails further tariff liberalization makes the North worse off and cannot be mutually beneficial. This logic extends beyond the case in which the export constraint is already absent. For tariff bounds in the intermediate range, the non-cooperative equilibrium features $\delta < 1$, so that eliminating the export constraint would increase global efficiency. However, if an agreement requires complete tariff elimination, these efficiency gains need not be sufficient to make cooperation mutually beneficial. Without transfers from the South to the North, the South cannot compensate the North for the loss of tariff revenue. Thus, in the baseline environment, even globally efficient deep liberalization may fail to be implementable as a mutually beneficial trade agreement.

Section 7 shows how this conclusion changes once removing the export constraint is costly for the South. With such costs, tariff reductions need not induce the South to eliminate its export constraint voluntarily. A shallow agreement may therefore reduce tariffs without restoring global efficiency. A mutually beneficial outcome may instead require a deep agreement that explicitly disciplines quantitative export restrictions and links tariff reductions to their removal. This mechanism provides a natural explanation for the widespread inclusion of such provisions in modern trade agreements (see Wu 2020, for further discussion).

6 A two-sector model

In our analysis so far, we have considered a single sector of production, which potentially uses two technologies. While a one-sector economy is particularly attractive for analytical tractability, it may overemphasize the role of the considered policy instruments and thus limit the practical relevance of our results. Therefore, we now extend our model to a two-sector economy, with a multi-stage production process for good Q_2 and a one-stage production process for an outside good Y. Each unit of the outside good requires the input of one unit of labor. To facilitate our analysis, we assume that good Y is consumed only in the North, where preferences are Cobb-Douglas and represented by the utility function $v_2 = \left(C_Q^2/L\right)^{\alpha} \left(C_Y^2/L\right)^{1-\alpha}$, with C_Q^2 and C_Y^2 denoting economy-wide Northern consumption of goods Q and Y, respectively, and $\alpha \in (0, 1)$ the Northern expenditure share on Q_2. Southern households consume C_Q^1

and C_Q^a of the homogeneous good Q. Since workers are homogeneous and the labor force in each country is fixed, per-capita consumption coincides with average welfare, given by $v_1 = C_Q^1/L + C_Q^a$, irrespective of labor mobility. We distinguish two cases in what follows. Subsection 6.1 assumes labor immobility and allocates $L/2$ workers to each sector, whereas Subsect. 6.2 analyzes the polar case of perfect inter-sectoral labor mobility.

6.1 A two-sector model with immobile labor

Under the assumption of labor immobility between the two sectors, the solution for the equilibrium production structure of Q_2 remains characterized by $\lambda_1 S_1 = -\ln[(x-1)/x]$ and the labor market clearing condition in Eq. (6′), with the important difference that x is now given by $2Q_0/(\delta\lambda_1 L)$. The change in the definition of x reflects the lower labor input of $L/2$ available for the production of good Q in the two-sector model without labor mobility. Moreover, open economy equilibrium wages and prices continue to be given by Eq. (8″). Restricting our analysis of the non-cooperative policy game to the same set of instruments as in the baseline model, namely $\tau_2 = \tau > 1$ and $\delta < 1$, and assuming that trade of good Y is not subject to any impediments,[13] we can use the market-clearing conditions for the two consumption goods to solve for the uniform wage paid by North and South in the outside sector. As formally shown in the Appendix, we compute:

$$w_Y = \frac{1-\alpha}{1+\alpha}\lambda_1 \left\{ \frac{(x-\Lambda)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} + \frac{\tau-1}{\tau}\frac{\delta(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} \right\}. \tag{15}$$

Noting the close resemblance between $\tilde{v}_2(\tau,\delta)$ in Eq. (12) and w_Y in Eq. (15), it is immediate that $dw_Y/d\tau > 0$ and $dw_Y/d\delta < 0$ (see the proof of Lemma 3). Moreover, we compute $C_Q^2/L = [\alpha/(1-\alpha)]w_Y$ and $C_Y^2/L = 1$, as a direct consequence of our assumption that Northern consumers have Cobb-Douglas preferences over the two consumption goods. This allows us to express welfare of the North (in per-capita terms) as $v_2 = [\alpha w_Y/(1-\alpha)]^\alpha$, implying that the properties of v_2 from Section 5 extend straightforwardly to the two-sector economy considered here.

Moreover, using $v_1 = C_Q^1/L + C_Q^a/L$, with $C_Q^1/L = (1/2)(\delta w_1 + w_Y)$ and $C_Q^a/L = (1/2)\lambda_1(1-\delta)/\{\exp[\lambda_1]-1\}$, we can express Southern welfare as follows:

$$v_1 = \frac{\lambda_1}{2}\left\{ [\delta(x-1)-\Lambda]\left[\frac{1-\alpha}{1+\alpha} + \frac{2\alpha}{1+\alpha}\frac{1}{\tau}\frac{\delta(x-1)}{\delta(x-1)+x-\Lambda}\right] + \frac{1-\delta}{\exp[\lambda_1]-1}\right\}. \tag{16}$$

Similar to Northern welfare, v_1 inherits important properties of the one-sector economy. For instance, welfare in the South decreases monotonically with τ. Additionally,

[13] Considering a homogeneous outside good that is freely tradable at zero costs is common practice in the literature on non-cooperative and cooperative trade policies (see Ossa 2011; Antràs and Staiger 2012a; Campolmi et al. 2014; Blanchard et al. 2026, for recent examples).

imposing a stricter export constraint δ reduces welfare through a reduction in labor income generated in global value chain production and thus lower imports of good Q from the North, while it increases local autarkic production output.

However, the impact of the two policy instruments on Southern welfare changes in the two-sector model. On the one hand, the negative consequences of a higher Northern tariff τ on Southern welfare are mitigated in the two-sector model. Higher tariff revenues increase Northern income and thus worldwide demand for good Y, resulting in a positive welfare effect on the South. On the other hand, a lower δ reduces income in the North, negatively affecting Southern income in sector Y. This counteracts the infra-marginal welfare gain from higher global value chain wages when lowering δ. As formally shown in Appendix A.5, these counteracting effects establish an interior solution of the non-cooperative policy game with $\delta > 0$ and $\tau = \overline{\tau}(\delta)$, showing that in the two-sector variant of our model with immobile labor, global value chain production is viable even if the two countries do not coordinate their trade policies. Despite these qualifications, the main insight from our one-sector model analysis that a trade agreement with $\tau = \delta = 1$ increases global value chain trade and can be beneficial for both countries continues to hold in the two-sector model. In particular, for high levels of α and thus a high Northern expenditure share for good Q, the results from the baseline model with a single production sector extend unchanged to the two-sector economy considered here.

6.2 A two-sector model with perfect labor mobility

With perfect inter-sectoral labor mobility, production in one sector may become unattractive, causing full specialization. Diversified production in the North requires $w_Y = w_2$. In the South, workers employed in the production of Q face ex ante uncertainty over wage realizations (w_1, w_1^a) and choose their occupations irreversibly (see Helpman and Itskhoki 2010; Egger and Kreickemeier 2012, for a similar approach). Under risk neutrality, diversification requires expected wages to be equalized $w_Y = \delta w_1 + (1-\delta) w_1^a$. Noting from our analysis in Sections 4 and 5, that $\lambda_1 > \lambda_2$ establishes $w_1^a \leq w_1 < w_2$, we can thus safely conclude that $w_Y < w_2$ so that the North specializes on the production of Q whereas the South either produces both goods or fully specializes on the production of the outside good.

With diversified production in the South, we can characterize the profit-maximizing production structure in the open economy by the upper threshold for the production stages conducted in the South, $\lambda_1 S_1 = -\ln[(x-1)/x]$, and the modified global labor market clearing condition in sector Q:

$$\bar{F}(x, \mu; \Lambda, \delta) \equiv -\ln\left(\frac{x-1}{x}\right) - \frac{1}{\Lambda}\ln\left(\frac{\delta\mu(x-1)-\Lambda}{\delta\mu(x-1)}\right) - \lambda_1 = 0, \qquad (17)$$

with $x = Q_0/(\lambda_1 \mu \delta L)$ as the initial input per unit of effective Southern labor used in global value chain production and with $\mu \in [0,1]$ as the share of Southern labor employed in the production of Q. It is easily confirmed that $\bar{F}'_x(\cdot) < 0$, whereas $\bar{F}'_\mu(\cdot) < 0$. This establishes $dx/d\mu < 0$, showing that changes in δ and changes in

μ exert similar effects on x and S_1, as both alter the share of Southern labor used in the global value chain production. However the interpretation of the two variables differs considerably. Whereas δ refers to an export quantity constraint and therefore to a trade policy instrument, the fraction of workers employed in the production of Q, μ, responds endogenously to changes in Northern and Southern trade policy due to inter-sectoral labor mobility.

The (residual) share of Southern labor used in the production of the outside good, $1-\mu$, can be pinned down in a next step by the market clearing condition for this good. Using the Cobb-Douglas property that Northern consumers spend a constant fraction $1-\alpha$ of their income on consumption of good Y, noting that per-capita Northern income is given by $w_2 + (\tau - 1)\lambda_1\delta\mu(x-1)p_1$, while per capita consumption of good Y equals $1-\mu$, and accounting for $w_Y = \delta w_1 + (1-\delta)w_1^a$ under diversified Southern production, we can express the market clearing condition for good Y as $(1-\mu)[\delta w_1 + (1-\delta)w_1^a] = (1-\alpha)[w_2 + (\tau-1)\lambda_1\delta\mu(x-1)p_1]$. We can further use the equation system outlined in Appendix A.2 — Eqs. (A.1) to (A.3) — to determine sector Q wages of the two countries as follows:

$$w_1 = \frac{\lambda_1}{\tau}\frac{(x-1)[\mu\delta(x-1)-\Lambda]}{\mu\delta(x-1)+x-\Lambda}, \qquad w_2 = \lambda_1\frac{(x-\Lambda)[\mu\delta(x-1)-\Lambda]}{\mu\delta(x-1)+x-\Lambda}. \tag{18}$$

Moreover, using Eq. (8″), the free-on-board price of intermediate good Q_1 is given by $p_1 = w_1/[\lambda_1(x-1)]$. We can thus rewrite the market clearing condition for good Y as

$$G(\mu, x; \Lambda, \delta, \tau) \equiv \frac{1-\mu}{1-\alpha}\left\{\delta + (1-\delta)\frac{\tau}{\exp[\lambda_1]-1}\frac{\mu\delta(x-1)+x-\Lambda}{(x-1)[\mu\delta(x-1)-\Lambda]}\right\} - \mu\delta(\tau-1) - \tau\frac{x-\Lambda}{x-1} = 0. \tag{19}$$

To ensure that $G(\mu, x; \Lambda, \delta, \tau) = 0$ has a solution in $\mu \in (0,1)$, we must consider τ to be sufficiently small. To see this, let us abstract from the export quantity constraint and set $\delta = 1$. Making use of $G(\mu, x; \Lambda, 1, \tau) \equiv \overline{G}(\mu, x; \Lambda, \tau)$, we find the fraction of Southern workers employed in the production of good Q to be implicitly determined by $\overline{G}(\mu, x; \Lambda, \tau) = [(1-\mu)/(1-\alpha)] - \mu(\tau-1) - \tau(x-\Lambda)/(x-1) = 0$. This implicit function is negatively sloped in μ, reaching a maximum of $\overline{G}(0, x; \Lambda, \tau) = 1/(1-\alpha) - \tau(x-\Lambda)/(x-1)$ at $\mu = 0$ and a minimum of $\overline{G}(1, x; \Lambda, \tau) = -\tau + 1 - \tau(x-\Lambda)/(x-1) < 0$ at $\mu = 1$. An interior solution with $\mu \in (0,1)$ therefore requires $1/(1-\alpha) > \tau(x-\Lambda)/(x-1)$, a condition which depends on endogenous x. Noting further that $\overline{G}(\mu, x; \Lambda, \tau)$ is positively sloped in x, we compute $dx/d\mu > 0$ from applying the implicit function theorem to $\overline{G}(\mu, x; \Lambda, \tau) = 0$.

Figure 4 depicts the two implicit functions from Eqs. (17) and (19) in the absence of an export quantity constraint ($\delta = 1$) and illustrates an interior solution in x and μ for the case of $\tau = 1$. With locus $\overline{F}(x, \mu; \Lambda, 1) = 0$ downward-sloping and locus $\overline{G}(x, \mu; \Lambda, \tau) = 0$ upward-sloping in the displayed (μ, x)-space, it is clear that the shown solution for x and μ must be unique. In the figure, we also show that an increase in the Northern import tariff shifts locus $\overline{G}(0, x; \Lambda, \tau) = 0$ leftwards,

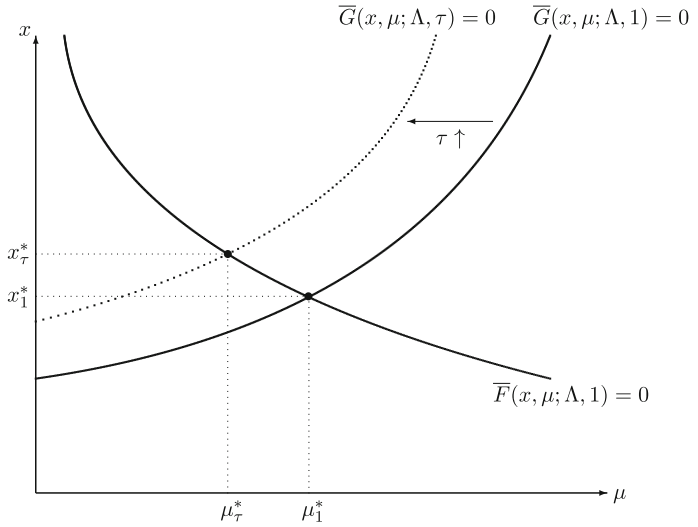


Fig. 4 An interior equilibrium in a two-sector model with labor mobility

inducing a decline in μ and an increase in x. The effects of increasing Northern tariffs is monotonic and it continues to hold up to a point in which $\tau = 1/(1-\alpha)$ establishes $\mu \to 0$ and $x \to \infty$. This shows that an interior solution of $\mu \in (0, 1)$ can be reached under the sufficient condition of $\tau < 1/(1-\alpha)$, making $1/(1-\alpha)$ an upper bound of Northern import tariffs that support diversified production in the South if $\delta = 1$.

For the case of a quantitative export constraint ($\delta < 1$), the analysis gets significantly more complicated, because for $\delta < 1$ the Northern import tariff has an additional upper constraint that is imposed by the outside option of workers in the autarkic production of good Q. From the condition $w_1 \geq w_1^a$, we derive an upper limit of τ supporting global value chain production in the South as $\{\exp[\lambda_1] - 1\}(x-1)[\mu\delta(x-1) - \Lambda]/[\mu\delta(x-1) + x - \Lambda] \equiv \overline{\tau}(\mu, x; \delta)$, with $\overline{\tau}(\mu, x; \delta) = \overline{\tau}$ if either $\mu = 0$ or $\delta = 0$. Condition $\tau \leq \overline{\tau}(\mu, x; \delta)$ therefore imposes an additional constraint in the derivation of endogenous $\{x, \mu\}$, which makes a general analytical solution infeasible.[14] However, we can determine the endogenous solutions of x and μ numerically and this is what we do when discussing optimal trade policies below.

To analyze optimal trade policies, we first determine Southern and Northern welfare per capita. Following the derivation steps in Subsect. 6.1, we compute per-capita Northern consumption of the two goods as $C_Q^2/L = [\alpha/(1-\alpha)](1-\mu)w_Y$ and $C_Y^2/L = (1-\mu)$. Plugging the latter into the Cobb-Douglas utility function gives $v_2 = (1-\mu)[\alpha w_Y/(1-\alpha)]^\alpha$. Similarly, we compute Southern per-capita consumption levels of $C_Q^1/L = \mu\delta w_1 + (1-\mu)w_Y$ and $C_Q^a/L = \mu\delta w_1^a$. Accounting for $w_Y = \delta w_1 + (1-\delta)w_1^a$, we obtain a welfare level equal to $v_1 = w_Y$. In the absence of an

[14] To see how the additional constraint impacts the derivation of x and μ, we can use the definition of $\overline{\tau}(\mu, x; \delta)$ to rewrite the implicit function in Eq. (19) as follows: $G(\mu, x; \Lambda, \delta, \tau) = [(1-\mu)/(1-\alpha)][\delta + (1-\delta)\tau/\overline{\tau}(\mu, x; \delta)] - \mu\delta(\tau - 1) - \tau(x - \Lambda)/(x-1) = 0$.

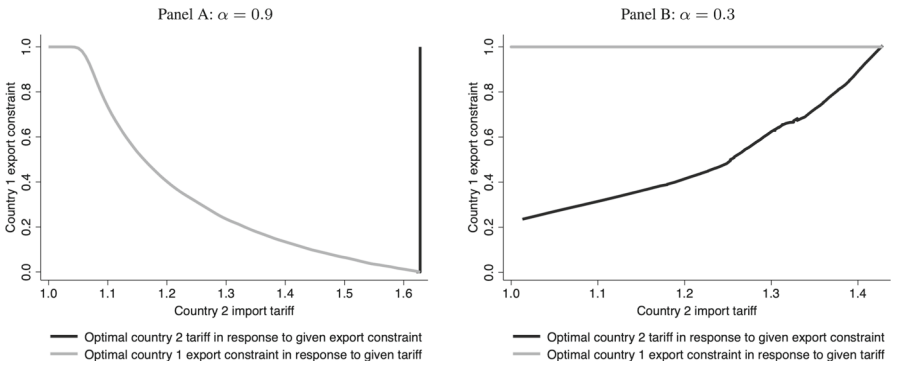


Fig. 5 Best-response policy choices of the two countries for $\lambda_1 = 0.95$ and $\lambda_2 = 0.05$

export quantity constraint ($\delta = 1$), a higher Northern tariff lowers Southern wages ceteris paribus, according to Eq. (18). In the two-sector model with labor mobility, this has two effects. On the one hand, it increases tax revenues of the North and, on the other hand, by lowering w_Y it improves the Northern terms of trade in consumption goods. Compared to the baseline model, this makes tariff increases more attractive for the North. However, higher tariffs also raise Northern demand for the outside good, which reduces μ and increases Southern wages, thereby worsening the North's terms of trade. These mechanisms are further complicated in environments with $\delta < 1$, so that the overall incentive of the North to increase τ to its maximum possible level is unclear in general.

Due to these complications, we explore optimal non-cooperative trade policies in a numerical exercise depicted in Figure 5, where we display the best (welfare-maximizing) responses of the South (δ) and the North (τ) for a given trade policy of the other economy. For our illustration, we consider significant technology differences and distinguish two scenarios: one with a high value of $\alpha = 0.9$ in Panel A and one with a lower value of $\alpha = 0.3$ in Panel B.

Panel A considers a high expenditure share for good Q_2. In this case, the North's incentive to appropriate rents along the value chain dominates the losses associated with a contraction of global value chain production. Hence, the North's best response is to set the import tariff on intermediates at the upper feasible bound, $\tau = \overline{\tau}$, for all values of the Southern export constraint; this is reflected by the vertical black locus in the figure. At such high tariff levels, the South's best response is to tighten its quantitative export constraint and set $\delta = 0$. The non-cooperative equilibrium therefore features $\mu = 0$ and $w_Y = w_1^a = \lambda_1/\{\exp[\lambda_1] - 1\}$.

Panel B considers a lower value of α, implying a larger expenditure share on the outside good in the North. In this case, a tariff increase generates weaker gains for the North than in Panel A, because the induced reallocation of Southern labor affects the price of good Y, which now carries greater weight in Northern expenditure. At the same time, tightening the export constraint becomes less attractive for the South. Although a lower δ raises the wage of workers employed in global value chain production, it lowers the average wage $\delta w_1 + (1-\delta) w_1^a$ and hence w_Y, reducing earnings for workers

employed in sector Y. In the illustrated case, the South's best response is therefore to set $\delta = 1$, thereby refraining from imposing a binding export constraint. The non-cooperative policy game nevertheless reaches a corner solution with $\tau = 1/(1-\alpha)$ and $\mu = 0$: global value chain production collapses because Southern workers allocate to the outside sector, not because the South imposes a binding export constraint.

In both scenarios illustrated in Figure 5, non-cooperative policy choices fail to implement the efficient open-economy allocation. A trade agreement with $\tau = \delta = 1$ restores global efficiency, stimulates global value chain trade, and (as illustrated in the next section) can be individually beneficial for both countries if changes in trade policies are costless. Therefore, the main insights from our previous analysis continue to hold in the model variant with two sectors and inter-sectoral labor mobility. This completes our discussion in this section.

7 Trade agreements if lifting the export constraints is costly

In Sections 5 and 6, we have demonstrated that a bilateral agreement to lower tariffs in the North and to lift Southern export constraints can foster mutually beneficial trade along the value chain. However, incentives for forming such an agreement exist in our model only if *(i)* the South imposes an export constraint in the non-cooperative scenario, and *(ii)* the gains from removing this constraint compensate the North for potential losses in tariff revenues. As highlighted by Proposition 3, both conditions are violated when initial tariffs are sufficiently low.

Our analysis thus far has overlooked, however, that changing existing export constraints may be costly, for instance, due to administrative or institutional expenditures associated with expanding authorized export volumes.[15] In this section, we look into such costs and, for simplicity, model them as an expenditure independent of the extent of δ-adjustment. We illustrate the outcome of this extended model in Figure 6, using a specific parameter configuration. Focusing on the two-sector model without labor mobility from Subsect. 6.1, we select a relatively high Northern expenditure share for good Q_2 ($\alpha = 0.9$) and consider significant productivity differences between the two economies, setting $\lambda_1 = 0.95$ and $\lambda_2 = 0.05$.[16]

In the left panel of Figure 6, we illustrate the welfare effects of a trade agreement imposing $\tau = \delta = 1$ for different levels of initial (non-cooperative) Northern import tariffs $\tau_{max} \leq \overline{\tau}$ and varying costs for lifting the export quantity constraint. These costs, denoted by f, are given in units of the consumption good Q_2 and are reported on a per-capita basis. To keep the discussion simple, we refer to f by the term *export licensing costs*. We identify four distinct outcomes:

1. **White Region**: This region represents combinations of initial tariffs (τ_{max}) and export licensing costs (f) that eliminate joint welfare gains from trade agreements. This region covers in particular cases with high levels of f.

[15] Such costs may include administrative, regulatory, or institutional expenses related to authorizing and monitoring higher export volumes, for example enhancing oversight capacity, processing additional export authorizations, or formalizing previously informal production activities.

[16] Since the results for the one-sector economy are similar to those for the two-sector model without labor mobility, we do not discuss them here.

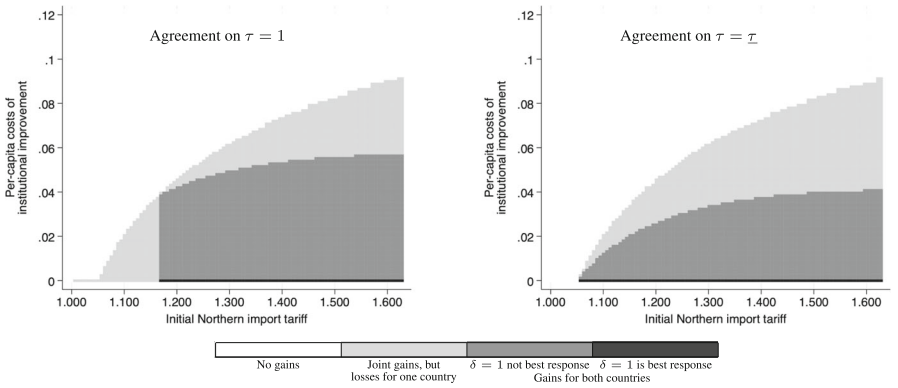


Fig. 6 Welfare effects of trade agreement (two sectors, $\lambda_1 = 0.95, \lambda_2 = 0.05, \alpha = 0.9$)

2. **Light Gray Region**: This region indicates combinations of τ_{max} and f where joint welfare gains exist, but one of the two countries incurs losses from signing a deep trade agreement. For instance, the North may experience losses if a significant reduction in tariff rates and a corresponding drop in tariff revenues are met with only a small reduction in Southern export constraints. According to our analysis in Section 5, this scenario is possible if τ_{max} is slightly higher than $\underline{\tau}$, while a small f prevents joint losses from the trade agreement. Moreover, the South may lose from a deep trade agreement if export licensing costs are considerable.
3. **Dark Gray Region**: This region indicates combinations of $\tau_{\max}$ and f for which both countries can be made better off by a trade agreement, but only if the agreement is deep. The reason is that the South would otherwise prefer to keep the export constraint in place. Successful implementation therefore requires an agreement that explicitly eliminates this constraint, rather than a shallow agreement that operates only through tariff reductions. The dark gray region thus identifies the parameter space in which deep agreements are relevant: a deep agreement that links tariff reductions to the removal of the export constraint can make both countries better off, whereas a shallow tariff agreement cannot implement the mutually beneficial outcome.
4. **Black Region**: This region represents scenarios in which both countries benefit from signing a trade agreement. Since setting $\delta = 1$ is the South's unilateral best response when the North eliminates its import tariff, such an agreement is attractive for both countries even without explicitly disciplining export constraints. A shallow agreement that operates only through tariff elimination is therefore sufficient to implement free trade. This outcome arises in particular when the costs of eliminating quantitative export constraints and expanding intermediate exports are negligible.

In the right panel of Figure 6, we consider the formation of a trade agreement with a tariff equal to $\underline{\tau} > 1$. We include this case for two reasons. First, preferential treatment of developing countries is one of the few exceptions to the most-favored-nation principle of the WTO, allowing for preferential trade agreements with positive

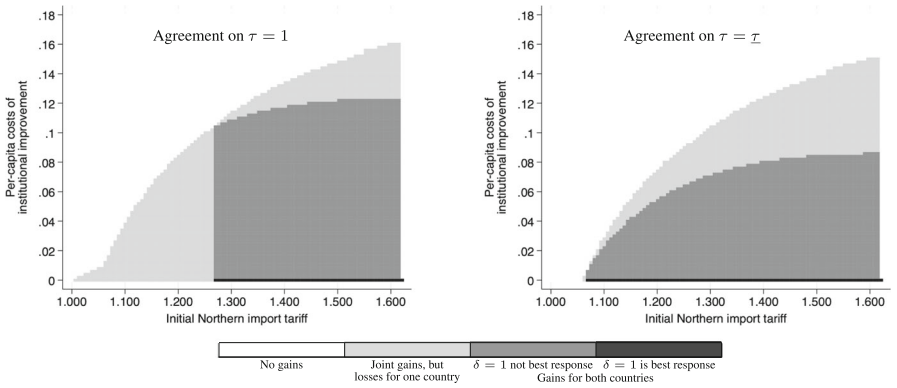


Fig. 7 Welfare effects of trade agreement if labor is mobile between sectors (two sectors, $\lambda_1 = 0.95$, $\lambda_2 = 0.05$, $\alpha = 0.9$)

tariff levels.[17] Second, previous research has established that setting zero tariffs is not necessarily welfare-optimal for signatories of a bilateral trade agreement (see Saggi et al. 2019, 2022). We find that positive tariffs make such agreements ceteris paribus more attractive for the North. However, the main insights from our previous analysis remain intact when extending our analysis to trade agreements with positive import tariffs.

We complete the discussion in this section by analyzing how perfect inter-sectoral labor mobility changes our insights from above. Starting from the theoretical model outlined in Subsect. 6.2, we illustrate in Figure 7 the scope for mutually beneficial trade agreements for this case. We use the same parameter values as in the left panel of Figure 5, acknowledging that a non-cooperative policy equilibrium of unconstrained governments would result in a corner solution with $\tau = \overline{\tau}$, $\delta = 0$, and $\mu = 0$. Using this as a starting point, we compute the welfare effects of a North-South trade agreement that implements $\tau = \delta = 1$ for initial tariffs $\tau_{max} \leq \overline{\tau}$. From the left panel of Figure 7, we observe that a tariff-only agreement is sufficient only in the black region, where setting $\delta = 1$ is the South's unilateral best response. Outside this region, tariff reductions alone do not implement the mutually beneficial outcome whenever the South would prefer to keep the export constraint in place. The dark gray region shows how this changes under a deep agreement: by explicitly requiring the removal of the export constraint, such an agreement can make both countries better off even for higher levels of f.

In the right panel, we consider an alternative agreement that lowers tariffs to $\tau = \underline{\tau}$ and establishes $\delta = 1$. The same logic applies. A shallow tariff agreement again fails to implement the mutually beneficial outcome whenever the South would keep the export constraint in place. By contrast, a deep agreement that links the tariff reduction to the removal of the Southern export constraint expands the parameter region in which both countries gain.

[17] Specifically, under the *enabling clause* of the WTO, a developed country can offer preferential treatment to imports from developing countries by unilaterally lowering its tariff.

8 Conclusions

This paper develops a two-country trade model with sequential production along the global value chain. In equilibrium, the value chain is segmented into two parts, with the technologically backward economy performing the initial production stages and the technologically advanced economy carrying out the later ones. The resulting trade pattern features exports of intermediates from the technologically backward economy in exchange for a homogeneous final good produced by the technologically advanced economy. Within this framework, tariff-based instruments are effective only for the technologically advanced economy, while the technologically backward economy can appropriate rents only through a quantity-based measure, such as a quantitative export constraint on intermediates. An import tariff on intermediates lowers wages in the technologically backward economy and increases welfare in the technologically advanced economy without affecting the production structure, making it equivalent to a lump-sum transfer. By contrast, a quantitative export constraint forces the technologically backward economy to rely partly on a less productive autarkic technology, reshaping the value chain and reducing specialization gains. This harms the technologically advanced economy, while the technologically backward economy may benefit when tariffs are high.

In the unconstrained non-cooperative policy game, the technologically advanced economy raises its import tariff up to the highest level consistent with the incentive-compatibility constraints of workers, while the technologically backward economy responds by refusing to export intermediates. This collapse of the global value chain drives both economies back to autarky. Starting from this benchmark, a trade agreement that eliminates tariffs raises welfare in both countries. However, after prior tariff liberalization has already restricted the technologically advanced economy's remaining tariff policy space, complete tariff elimination need not be mutually beneficial. If the technologically advanced economy faces a sufficiently low upper bound on its non-cooperative tariff, the technologically backward economy will unilaterally remove its export constraint even without additional tariff concessions, leaving the technologically advanced economy with no incentive to cut tariffs further. Likewise, if eliminating the quantitative export constraint is costly for the technologically backward economy, it may prefer to keep the constraint in place despite tariff reductions. In such cases, a shallow agreement focusing solely on tariff elimination is not attractive to one or both countries. A mutually beneficial agreement may then require a *deep* trade agreement that disciplines quantitative export restrictions and links tariff reductions to their removal. Extensions of the model to include two production sectors do not alter these core insights.

Taken together, our theoretical analysis helps explain two empirical patterns: the growing incidence of quantitative export restrictions during a period of rapid expansion in global value chain trade, and the prevalence of deep trade agreements containing provisions disciplining such restrictions. Our central contribution is to show that countries differ fundamentally in the effectiveness of tariff- versus quantity-based instruments for rent appropriation along the value chain, and that this asymmetry creates a rationale for deep trade agreements. At the same time, the sequential production framework we employ is deliberately parsimonious. It abstracts from important real-world features,

such as multiple countries, multiple sectors, and the coexistence of shallow and deep agreements. Extending the analysis to a richer environment would be a promising direction for future theoretical and empirical research on trade policy in global value chains.

A Theoretical appendix

A.1 Derivation of equation system (7)

From Eq. (2), we know that $[p(s+ds) - p(s)]/ds = [p(s)\lambda_c + w_c]/(1 - \lambda_c ds)$ holds whenever $Q_c(s') > 0$ for all $s' \in (s, s+ds]$. Taking the limit of $ds \to 0$ gives the differential equation $p'(s) = p(s)\lambda_c + w_c$. We can now solve this differential equation for the general solution $p(s) = -w_c/\lambda_c + B_c \exp[\lambda_c s]$. Making use of the equilibrium production structure from Lemma 1, we compute $p(0) = -w_1/\lambda_1 + B_1 = 0$ and thus $B_1 = w_1/\lambda_1$. This establishes $p(s) = \{\exp[\lambda_1 s] - 1\} w_1/\lambda_1$ for all $s \in (0, S_1]$ and thus the first line in system (7) when setting $s = S_1$ and $p(S_1) = p_1$. Moreover, evaluating $p(s) = -w_2/\lambda_2 + B_2 \exp[\lambda_2 s]$ at $s = S_1$, we compute $B_2 = [p_1 + w_2/\lambda_2] \exp[-\lambda_2 S_1]$, which establishes $p(s) = p_1 \exp[\lambda_2(s - S_1)] + (w_2/\lambda_2)\{\exp[\lambda_2(s - S_1)] - 1\}$. Evaluated at $s = 1$ and $p(1) = 1$, this gives the second line of equation system (7).

In a final step, we acknowledge that a profit-maximizing choice of executing stages $s > S_1$ in country 2 and executing stages $s < S_1$ in country 1 requires

$$p(S_1 + ds) - \frac{p(S_1) + w_2 ds}{1 - \lambda_2 ds} \geq p(S_1 + ds) - \frac{p(S_1) + w_1 ds}{1 - \lambda_1 ds},$$
$$p(S_1) - \frac{p(S_1 - ds) + w_1 ds}{1 - \lambda_1 ds} \geq p(S_1) - \frac{p(S_1 - ds) + w_2 ds}{1 - \lambda_2 ds}$$

to simultaneously hold according to Eq. (2). Making use of standard mathematical manipulation and setting $p(S_1) = p_1$, we can reformulate the two conditions for the limiting case of $ds \to 0$ as

$$p_1(\lambda_1 - \lambda_2) \geq w_2 - w_1, \qquad w_2 - w_1 \geq p_1(\lambda_1 - \lambda_2),$$

which jointly establish the third line of equation system (7). This completes the proof.

A.2 Derivation of w_1, w_2, and p_1 in Eqs. (8′) and (8″)

The derivation steps have the same order as in Appendix A.1. Thereby, we allow for ad-valorem export tariffs on intermediates by country 1, denoted $\tau_1 > 0$, and for ad-valorem import tariffs of intermediates by country 2, denoted $\tau_2 > 0$. Moreover, we consider export constraints for intermediates imposed by country 1, which are captured by $\delta < 1$. Tariffs on the final good are irrelevant for the subsequent derivation.

Starting point is differential equation $p_c'(s) = p_c(s)\lambda_c + w_c$, where a country index is attached to the price, because $\delta < 1$ drives a wedge between the price per

produced and the price per shipped unit and because $\tau_1, \tau_2 > 1$ drive a wedge between the export and the import price. The differential equation has the general solution $p_c(s) = -w_c/\lambda_c + B_c \exp[\lambda_c s]$. In view of $p_1(0) = -w_1/\lambda_1 + B_1 = 0$, we get $B_1 = w_1/\lambda_1$. This establishes $p_1(s) = \{\exp[\lambda_1 s] - 1\} w_1/\lambda_1$ for all $s \in (0, S_1]$. Setting $p_1(S_1) = p_1$, we obtain for $s = S_1$:

$$p_1 = \{\exp[\lambda_1 S_1] - 1\} \frac{w_1}{\lambda_1}. \tag{A.1}$$

Moreover, evaluating $p_2(s) = -w_2/\lambda_2 + B_2 \exp[\lambda_2 s]$ at $s = S_1$, we compute, using $p_2(S_1) = \tau_1 \tau_2 p_1$, $B_2 = [\tau_1 \tau_2 p_1 + w_2/\lambda_2] \exp[-\lambda_2 S_1]$. This establishes $p_2(s) = \tau_1 \tau_2 p_1 \exp[\lambda_2 (s - S_1)] + (w_2/\lambda_2)\{\exp[\lambda_2 (s - S_1)] - 1\}$, which evaluated at $s = 1$ and setting $p_2(1) = 1$ gives

$$1 = \exp[\lambda_2 (1 - S_1)] \tau_1 \tau_2 p_1 + \{\exp[\lambda_2 (1 - S_1)] - 1\} \frac{w_2}{\lambda_2}. \tag{A.2}$$

Finally, we acknowledge that a profit-maximizing choice of executing stages $s > S_1$ in country 2 and executing stages $s < S_1$ in country 1 requires

$$p_2(S_1 + ds) - \frac{p_2(S_1) + w_2 ds}{1 - \lambda_2 ds} \geq p_2(S_1 + ds) - \tau_1 \tau_2 \frac{p_1(S_1) + w_1 ds}{1 - \lambda_1 ds},$$
$$p_2(S_1) - \tau_1 \tau_2 \frac{p_1(S_1 - ds) + w_1 ds}{1 - \lambda_1 ds} \geq p_2(S_1) - \frac{p_2(S_1 - ds) + w_2 ds}{1 - \lambda_2 ds}$$

to simultaneously hold according to Eq. (2). Making use of standard mathematical manipulation and setting $p_1(S_1) = p_1$, $p_2(S_1) = \tau_1 \tau_2 p_1$, we can reformulate the two conditions for the limiting case of $ds \to 0$ as follows

$$\tau_1 \tau_2 p_1 (\lambda_1 - \lambda_2) \geq w_2 - \tau_1 \tau_2 w_1, \qquad w_2 - \tau_1 \tau_2 w_1 \geq \tau_1 \tau_2 p_1 (\lambda_1 - \lambda_2),$$

which jointly establish

$$w_2 = \tau_1 \tau_2 w_1 + \tau_1 \tau_2 p_1 (\lambda_1 - \lambda_2). \tag{A.3}$$

We can combine Eqs. (A.1)-(A.3) and make use of $\lambda_1 S_1 = \ln[(x - 1)/x]$ and $\lambda_2 (1 - S_1) = \ln\left[\{\delta(x - 1) - \Lambda\}/\{\delta(x - 1)\}\right]$ to determine the explicit solutions for w_1, w_2 and p_1 in Eqs. (8′) and (8″), where $\delta = 1$ is assumed in the former. This completes the proof.

A.3 Proof of Lemma 3

Noting $dx/d\tau = 0$ from Eq. (6′), $\partial \tilde{v}_1(\cdot)/\partial \tau < 0$ and $\partial \tilde{v}_2(\cdot)/\partial \tau > 0$ directly follow from Eqs. (11) and (12), respectively. Moreover, applying the implicit function theorem

to Eq. (6′), we obtain

$$\frac{dx}{d\delta} = -\frac{x}{\delta}\frac{x-1}{\delta(x-1)+x-\Lambda} < 0, \quad \frac{d[\delta(x-1)]}{d\delta} = \frac{(x-1)[\delta(x-1)-\Lambda]}{\delta(x-1)+x-\Lambda} > 0. \tag{A.4}$$

Then, differentiating w_1 from Eq. (8″) with respect to δ gives $\partial w_1/\partial\delta = -(w_1/\delta)x(1-\Lambda)[\delta(x-1)+x-\Lambda]^{-2} < 0$.[18] Noting $\tilde{v}_1(\tau,\delta) = \delta w_1 + (1-\delta)w_1^a$ from Eq. (11), we therefore compute

$$\frac{\partial\tilde{v}_1(\cdot)}{\partial\delta} = w_1\left[1 - \frac{x(1-\Lambda)}{[\delta(x-1)+x-\Lambda]^2}\right] - \frac{\lambda_1}{\exp[\lambda_1]-1}, \tag{A.5}$$

while the second partial derivative is given by

$$\frac{\partial^2\tilde{v}_1(\cdot)}{\partial\delta^2} = -\frac{w_1}{\delta}\frac{x(1-\Lambda)}{[\delta(x-1)+x-\Lambda]^2}\left\{\frac{\delta(x-1)}{\delta(x-1)+x-\Lambda} + \frac{(1-\Lambda)[\delta(x-1)-\Lambda]}{[\delta(x-1)+x-\Lambda]^2} + 2\frac{(x-1)[x+\delta\Lambda-\delta^2(x-1)]}{[\delta(x-1)+x-\Lambda]^2}\right\} < 0.$$

This implies that if $\tilde{v}_1(\cdot)$ has an extremum in δ on the unit interval, it must be a maximum.

Acknowledging $\partial^2\tilde{v}_1(\cdot)/\partial\delta\partial\tau < 0$ and

$$\lim_{\delta\to 0}\frac{\partial\tilde{v}_1(\cdot)}{\partial\delta} = \frac{1}{\tau}\frac{\lambda_2}{\exp[\lambda_2]-1} - \frac{\lambda_1}{\exp[\lambda_1]-1}, \tag{A.6}$$

we can determine a critical $\overline{\tau}$ by setting $\partial\tilde{v}_1(\overline{\tau},0)/\partial\delta = 0$.[19] It is then immediate that $\partial\tilde{v}_1(\tau,0)/\partial\delta >, =, < 0$ if $\overline{\tau} >, =, < \tau$. In view of $\partial^2\tilde{v}_1(\cdot)/\partial\delta^2 < 0$ it follows that the welfare maximizing level of δ is larger than zero if $\tau < \overline{\tau}$. We can compute a second critical level of τ, $\underline{\tau}$, by setting $\partial\tilde{v}_1(\underline{\tau},1)/\partial\delta = 0$. Due to $\partial^2\tilde{v}_1(\cdot)/\partial\delta\partial\tau < 0$, it is then immediate that $\partial\tilde{v}_1(\tau,1)/\partial\delta >, =, < 0$ if $\underline{\tau} >, =, < \tau$. In view of $\partial^2\tilde{v}_1(\cdot)/\partial\delta^2 < 0$, it follows that the welfare maximizing level of δ is smaller than one if $\tau > \underline{\tau}$. Setting $\partial\tilde{v}_1/d\delta = 0$ determines τ as implicit function of δ, which we denote by $\hat{\tau}(\delta)$. Making use of $\partial^2\tilde{v}_1/\partial\delta^2 < 0$, $\partial^2\tilde{v}_1/\partial\delta\partial\tau < 0$, we have $\hat{\tau}'(\delta) < 0$, which establishes $\underline{\tau} < \overline{\tau}$. Moreover, we can show that $\underline{\tau} > 1$. Since the formal proof for this result is tedious, we have deferred it to the Online Appendix to this manuscript.

We now turn to the impact of changes in parameter δ on welfare in the North. For this purpose, we first differentiate w_2 from Eq. (8″) with respect to δ. Making use of the derivatives in Eq. (A.4), we compute $\partial w_2/\partial\delta = \tau w_1 x(1-\Lambda)[\delta(x-1)+x-\Lambda]^{-2} > 0$ (where $\tau_1 = 1$ and $\tau_2 = \tau$ have been used). Accounting for $\tilde{v}_2(\tau,\delta) = w_2 + (\tau -$

[18] Similarly, we compute $\partial p_1/\partial\delta = (p_1/\delta)x(x-1)(1+\delta)[\delta(x-1)+x-\Lambda]^{-2} > 0$.

[19] From Eq. (A.6), we can compute the explicit solution $\overline{\tau} = \frac{\lambda_2}{\lambda_1}\frac{\exp[\lambda_1]-1}{\exp[\lambda_2]-1} > 1$, which corresponds to $\lim_{\delta\to 0}\overline{\tau}(\delta)$ in Eq. (13).

$1)\lambda_1 p_1 \delta(x-1) = w_2 + (\tau-1)\delta w_1$ from Eqs. (8″) and (12), we further compute

$$\frac{\partial \tilde{v}_2(\cdot)}{\partial \delta} = \tau w_1 \frac{x(1-\Lambda)}{[\delta(x-1)+x-\Lambda]^2} + (\tau-1)w_1\left[1 - \frac{x(1-\Lambda)}{[\delta(x-1)+x-\Lambda]^2}\right] > 0, \tag{A.7}$$

which completes the proof.

A.4 Proof of Proposition 2

We consider $\tau = \delta = 1$. Then, setting $\Lambda = 1$ establishes $x = 2\exp[\lambda_1]/\{\exp[\lambda_1]-1\}$, according to Footnote 11. In this case, we have $\tilde{v}_1(1,1) = \lambda_1/\{\exp[\lambda_1]-1\}$ and thus $\tilde{v}_1(1,1) = v_1^{nc}$, according to Eqs. (11) and (14). Totally differentiating $\tilde{v}_1(1,1)$ with respect to λ_2 further establishes

$$\frac{d\tilde{v}_1(1,1)}{d\lambda_2} = \lambda_1 \frac{2x(x-1-\Lambda)+\Lambda(1+\Lambda)}{(2x-1-\Lambda)^2}\frac{dx}{d\Lambda}\frac{d\Lambda}{d\lambda_2} - \lambda_1 \frac{x(x-1)}{(2x-1-\Lambda)^2}\frac{d\Lambda}{d\lambda_2}.$$

Applying the implicit function theorem to Eq. (6′) and evaluating the resulting expression at $\delta = 1$, we obtain

$$\frac{dx}{d\Lambda} = \frac{x}{\Lambda}\frac{x-1}{2x-1-\Lambda}\left[1 + \frac{x-1-\Lambda}{\Lambda}\ln\left(\frac{x-1-\Lambda}{\Lambda}\right)\right] > 0. \tag{A.8}$$

Using the derivative in Eq. (A.8) together with $d\Lambda/d\lambda_2 = 1/\lambda_1$, we compute

$$\frac{d\tilde{v}_1(1,1)}{d\lambda_2} = \frac{x(x-1)}{(2x-1-\Lambda)^2}\left\{\frac{2x(x-1-\Lambda)+\Lambda(1+\Lambda)}{\Lambda(2x-1-\Lambda)} \times \left[1 + \frac{x-1-\Lambda}{\Lambda}\ln\left(\frac{x-1-\Lambda}{x-1}\right)\right] - 1\right\}.$$

From the properties of the natural logarithm it is well known that, for $y > 0$, we have $\ln y \le y - 1$. We next show that $d\tilde{v}_1(1,1)/d\lambda_2 < 0$.

Noting that $x(x-1)/(2x-1-\Lambda)^2 > 0$, the sign of $d\tilde{v}_1(1,1)/d\lambda_2$ is determined by the sign of $A(x,\Lambda)B(x,\Lambda) - 1$, where

$$A(x,\Lambda) \equiv \frac{2x(x-1-\Lambda)+\Lambda(1+\Lambda)}{\Lambda(2x-1-\Lambda)}, \quad \text{and}$$
$$B(x,\Lambda) \equiv 1 + \frac{x-1-\Lambda}{\Lambda}\ln\left(\frac{x-1-\Lambda}{x-1}\right).$$

We account for $x > 1+\Lambda$ and make use of the auxiliary term $y \equiv \Lambda/(x-1) \in (0,1)$. This implies that $B(x,\Lambda) = 1+[(1-y)/y]\ln(1-y)$. Noting further that $\ln(1-y) <$

$-y - y^2/2$,[20] we compute $B(x, \Lambda) < y(1+y)/2 = \Lambda(x-1-\Lambda)/[2(x-1)^2]$ and thus

$$A(x, \Lambda)B(x, \Lambda) - 1 < \frac{[2x(x-1-\Lambda) + \Lambda(1+\Lambda)](x-1+\Lambda)}{2(2x-1-\Lambda)(x-1)^2} - 1 < 0.$$

This establishes $d\tilde{v}_1(1,1)/d\lambda_2 < 0$, which is sufficient for $\tilde{v}_1(1,1) > v_1^{nc}$ to hold for all $\Lambda < 1$ and thus $\lambda_2 < \lambda_1$.

To determine the ranking of $\tilde{v}_2(1,1)$ and v_2^{nc}, we first rewrite Eq. (6) as follows

$$\hat{F}(x, \Lambda) \equiv -\Lambda \ln\left(\frac{x-1}{x}\right) - \ln\left(\frac{x-1-\Lambda}{x-1}\right) - \lambda_2 = 0. \tag{A.9}$$

Applying the implicit function theorem, we compute

$$\frac{dx}{d\Lambda} = \frac{x}{\Lambda}\frac{x-1}{2x-1-\Lambda}\left[1 - (x-1-\Lambda)\ln\left(\frac{x-1}{x}\right)\right] > 0. \tag{A.10}$$

For a given mistake rate in the North (i.e. for a given level of λ_2), x increases in Λ. Setting $\Lambda = 1$ gives $\lambda_1 = \lambda_2$, $x = 2\exp[\lambda_2]/\{\exp[\lambda_2] - 1\}$, and $\tilde{v}_2(1,1) = \lambda_2/\{\exp[\lambda_2] - 1\}$. This establishes $\tilde{v}_2(1,1) = v_2^{nc}$. Totally differentiating $\tilde{v}_2(1,1)$ with respect to λ_1 gives

$$\begin{aligned}\frac{d\tilde{v}_2(1,1)}{d\lambda_1} = {} & \lambda_1 \frac{2x(x-1-\Lambda)+1+\Lambda}{(2x-1-\Lambda)^2}\frac{dx}{d\Lambda}\frac{d\Lambda}{d\lambda_1} \\ & - \lambda_1 \frac{(2x-1-\Lambda)(x-1-\Lambda) + x(x-\Lambda)}{(2x-1-\Lambda)^2}\frac{d\Lambda}{d\lambda_1} \\ & + \frac{(x-\Lambda)(x-1-\Lambda)}{2x-1-\Lambda}.\end{aligned}$$

Substituting $dx/d\Lambda$ from Eq. (A.10) and $d\Lambda/d\lambda_1 = -\Lambda/\lambda_1$, we compute

$$\begin{aligned}\frac{d\tilde{v}_2(1,1)}{d\lambda_1} = {} & -\frac{2x(x-1-\Lambda)+1+\Lambda}{(2x-1-\Lambda)^2} \\ & \times \frac{x(x-1)}{2x-1-\Lambda}\left[1 - (x-1-\Lambda)\ln\left(\frac{x-1}{x}\right)\right] \\ & + \frac{\Lambda x(x-\Lambda)}{(2x-1-\Lambda)^2} + \frac{x(x-1-\Lambda)}{2x-1-\Lambda}.\end{aligned}$$

Rearranging terms, we obtain $d\tilde{v}_2(1,1)/d\lambda_1 = x(x-1)(x-1-\Lambda)[2x(x-1-\Lambda)+1+\Lambda](2x-1-\Lambda)^{-3}\hat{f}(x,\lambda)$, with

$$\hat{f}(x, \Lambda) = \frac{2(x-\Lambda)}{2x(x-1-\Lambda)+1+\Lambda} + \ln\left(\frac{x-1}{x}\right).$$

[20] To see this, we define $h(y) := \ln(1-y) + y + y^2/2$, with $h(0) = 0$ and $h'(y) = -y^2/(1-y)$, which is lower than zero for all $y \in (0,1)$. This establishes $h(y) < 0$ for all admissible y.

Partially differentiating $\hat{f}(x, \Lambda)$ with respect to Λ gives $\hat{f}'_{\Lambda}(x, \Lambda) = 2(x-1)[2x(x-1-\Lambda)+1+\Lambda]^{-2} > 0$ and thus $\hat{f}(x, \Lambda) > \hat{f}(x, 0)$. Moreover, differentiating $\hat{f}(x, 0) = 2x/[2x(x-1)+1] + \ln[(x-1)/x]$ with respect to x, we obtain

$$\hat{f}'_x(x, 0) = -\frac{(4x-6)x(x-1)-1}{x(x-1)[2x(x-1)+1]^2}.$$

Since $\hat{f}'_x(x, 0) < 0$ holds for all $x > 2$ and since $x > 2$ is imposed by Footnote 11, we can safely conclude that $\hat{f}(x, \Lambda) > \hat{f}(x, 0) > \lim_{x\to\infty} \hat{f}(x, 0) = 0$. This is sufficient for $d\tilde{v}_2(1, 1)/d\lambda_1 > 0$ and thus for $\tilde{v}_2(1, 1) > v_2^{nc}$ to hold for all $\Lambda < 1$, i.e. for all $\lambda_1 > \lambda_2$, which completes the proof.

A.5 Derivation details for the analysis in Subsect. 6.1

We first note that total income in country 2 is given by $(w_Y + w_2)L/2 + (\tau - 1)p_1 Q_1$. Moreover, the assumption of Cobb-Douglas preferences in the North implies that consumers spend constant fractions α and $(1-\alpha)$ of their expenditures on the consumption of good Q, denoted by C_Q^2, and the consumption of good Y, denoted by C_Y^2, respectively. Northern consumption of the two goods then follows as $C_Q^2 = \alpha[(w_2 + w_Y)L/2 + (\tau-1)p_1 Q_1]$, $p_Y C_Y^2 = (1-\alpha)[(w_2 + w_Y)L/2 + (\tau-1)p_1 Q_1]$, where $p_Y = w_Y$ is the consumer price of good Y. Noting further that consumption of good Y only exists in county 2, we have $C_Y^2 = L$, while $Q_1 = Q_0 - \delta\lambda_1 L/2 = \delta\lambda_1(x-1)L/2$ follows from our technology assumption. Sector Y wages are then given by $w_Y = [(1-\alpha)/(1+\alpha)][w_2 + (\tau-1)\lambda_1 p_1 \delta(x-1)]$. Substituting for w_2 and p_1 from the unchanged Eq. (8″), we can solve for wage rate w_Y in Eq. (15).

Moreover, Northern consumption of good Q can be derived, using our previous insight that households devote constant shares of their expenditures on the consumption of the two goods. This establishes $C_Q^2 = [\alpha/(1-\alpha)]w_Y L$. Northern welfare can then be obtained by substituting $C_Y^2/L = 1$ and $C_Q^2/L = [\alpha/(1-\alpha)]w_Y$ into the utility function $u_2 = (C_Q^2/L)^{\alpha}(C_Y/L)^{1-\alpha}$. This establishes household welfare $v_2 = [\alpha w_Y/(1-\alpha)]^{\alpha}$. To determine the impact of changes in τ and δ on v_2, we need to determine their impact on w_Y. Introducing the auxiliary expression $X_1(\delta) \equiv \delta(x-1)$ and $X_2(\delta) = x - \Lambda$, we rewrite Eq. (15) as follows

$$w_Y = \lambda_1 \frac{1-\alpha}{1+\alpha}[X_1(\delta) - \Lambda]\left[1 - \frac{1}{\tau}\frac{X_1(\delta)}{X_1(\delta) + X_2(\delta)}\right]. \tag{A.11}$$

Acknowledging $dX_1(\delta)/d\tau = dX_2(\delta)/d\tau = 0$ from Eq. (6′), we obtain $dw_Y/d\tau > 0$. Moreover, making use of Eq. (A.4), we compute

$$X_1'(\delta) = \frac{1}{\delta}\frac{X_1(\delta)[X_1(\delta) - \Lambda]}{X_1(\delta) + X_2(\delta)} > 0, \qquad X_2'(\delta) = -\frac{1}{\delta}\frac{X_2(\delta) + \Lambda}{X_1(\delta) - \Lambda}X_1'(\delta) < 0. \tag{A.12}$$

This allows us to determine

$$\frac{dw_Y}{d\delta} = \lambda_1 \frac{1-\alpha}{1+\alpha} X_1'(\delta) \left\{ 1 - \frac{1}{\tau} \left[1 - \frac{X_2(\delta)+\Lambda}{X_1(\delta)+X_2(\delta)} \frac{1-\Lambda}{X_1(\delta)+X_2(\delta)} \right] \right\} > 0.$$

Putting together, we can safely conclude that $dv_2/d\tau > 0$, implying that $\tau = \overline{\tau}(\delta)$ remains to be the unilaterally preferred policy choice of the North. At the same time $dv_2/d\delta > 0$ implies that the North benefits from a lower Southern export constraint.

Let us now turn to the South, where the balanced budget constraint establishes, in analogy to Eq. (9′), $C_Q^1 = (w_Y + \delta w_1)L/2$ and $C_Q^a = Q^a$. Making use of $Q^a = (1-\delta)\lambda_1(L/2)/\{\exp[\lambda_1] - 1\}$, substituting w_1 from Eq. (8″), and acknowledging $v_1 = C_Q^1/L + C_Q^a/L$, we can solve for Southern welfare in Eq. (16). Using auxiliary variables $X_1(\delta)$, $X_2(\delta)$, we can express v_1 in a more compact form as

$$v_1 = \frac{\lambda_1}{2} \left\{ [X_1(\delta) - \Lambda] \left[\frac{1-\alpha}{1+\alpha} + \frac{2\alpha}{1+\alpha} \frac{1}{\tau} \frac{X_1(\delta)}{X_1(\delta)+X_2(\delta)} \right] + \frac{1-\delta}{\exp[\lambda_1]-1} \right\}, \tag{A.13}$$

with $dv_1/d\tau < 0$ following from $dX_1(\delta)/d\tau = dX_2(\delta)/d\tau = 0$. Differentiating v_1 with respect to δ gives

$$\frac{\partial v_1}{\partial \delta} = \frac{\lambda_1}{2} \left\{ X_1'(\delta) \left[\frac{1-\alpha}{1+\alpha} + \frac{2\alpha}{1+\alpha} \frac{1}{\tau} \left(1 - \frac{X_2(\delta)+\Lambda}{X_1(\delta)+X_2(\delta)} \frac{1-\Lambda}{X_1(\delta)+X_2(\delta)} \right) \right] - \frac{1}{\exp[\lambda_1]-1} \right\}.$$

We now evaluate $\partial v_1/\partial\delta$ at the dominant policy strategy of the North: $\tau = \overline{\tau}(\delta) = \{\exp[\lambda_1] - 1\} X_1'(\delta)$. This gives

$$\left.\frac{\partial v_1}{\partial \delta}\right|_{\tau=\overline{\tau}(\delta)} = \frac{\lambda_1}{2} \left\{ \frac{1-\alpha}{1+\alpha} \left[X_1'(\delta) - \frac{1}{\exp[\lambda_1]-1} \right] - \frac{2\alpha}{1+\alpha} \times \frac{1}{\exp[\lambda_1]-1} \frac{X_2(\delta)+\Lambda}{X_1(\delta)+X_2(\delta)} \frac{1-\Lambda}{X_1(\delta)+X_2(\delta)} \right\}.$$

We can next determine $\partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)}$ at the two limiting cases of $\delta \to 0$ and $\delta \to 1$. We begin with the former and note that $\lim_{\delta\to 0} X_1'(\delta) = \lim_{\delta\to 0}[X_1(\delta) - \Lambda]$, according to the definition of $X_1(\delta)$, $X_2(\delta)$ and Eq. (A.12). Making use of Eq. (6′), we further compute $\lim_{\delta\to 0}[X_1(\delta) - \Lambda] = \Lambda/\{\exp[\lambda_2] - 1\}$. Finally, noting $\lim_{\delta\to 0}[X_2(\delta) + \Lambda]/[X_1(\delta) + X_2(\delta)]^2 = 0$, it follows that

$$\lim_{\delta\to 0} \left.\frac{\partial v_1}{\partial \delta}\right|_{\tau=\overline{\tau}(\delta)} = \frac{1}{2} \frac{1-\alpha}{1+\alpha} \left[\frac{\lambda_2}{\exp[\lambda_2]-1} - \frac{\lambda_1}{\exp[\lambda_1]-1} \right] > 0,$$

where the positive sign can be verified by noting that $\exp[z] - 1 = \int_0^z \exp[t]dt = z\int_0^1 \exp[sz]ds$. This establishes $f(z) \equiv z/\{\exp[z] - 1\} = \{\int_0^1 \exp[sz]ds\}^{-1}$ and

therefore $f'(z) < 0$. Hence, the positive sign of $\lim_{\delta\to 0} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)}$ follows from $\lambda_1 > \lambda_2$. Summing up, we conclude that if the North sets $\tau = \overline{\tau}(0)$, the best-response of the South is to set $\delta > 0$.

Considering the limiting case of $\delta \to 1$, we can note that $\lim_{\delta\to 1} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)} < 0$ must hold for α sufficiently close to 1. Moreover, $\lim_{\delta\to 1} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)} < 0$ extends to arbitrary levels of α if $\{\exp[\lambda_1]-1\}^{-1} > X_1'(1) = (x-1)(x-1-\Lambda)/(2x-1-\Lambda) \equiv a(\Lambda)$, where Eq. (6) implicitly determines x as a function of Λ for $\delta = 1$, with $dx/d\Lambda > 0$, according to Eq. (A.8). We thus compute

$$a'(\Lambda) = \frac{x(x-1)}{(2x-1-\Lambda)^2}\left[\rho(\Lambda)\ln\rho(\Lambda) + \frac{2(x-\Lambda)(x-1-\Lambda)}{\Lambda(2x-1-\Lambda)}[1+\rho(\Lambda)\ln\rho(\Lambda)]\right\}, \tag{A.14}$$

with $\rho(\Lambda) \equiv (x-1-\Lambda)/\Lambda$.

We can further note that $\rho(\Lambda)\ln\rho(\Lambda)$ has a unique minimum equal to $-\exp[-1]$ at $\rho(\Lambda) = \exp[-1]$. This implies that $1+\rho(\Lambda)\ln\rho(\Lambda) \geq 1-\exp[-1] > 0$. Moreover, we find that $2(x-\Lambda)(x-1-\Lambda)/[\Lambda(2x-1-\Lambda)] = [(2x-1-\Lambda)+(1-\Lambda)](x-1-\Lambda)/[\Lambda(2x-1-\Lambda)] > (x-2)$, where the inequality follows from $x-1-\Lambda > x-2$ and $\Lambda \in (0,1)$. This implies that

$$\begin{aligned}&\rho(\Lambda)\ln\rho(\Lambda) + \frac{2(x-\Lambda)(x-\Lambda-1)}{\Lambda(2x-1-\Lambda)}[1+\rho(\Lambda)\ln\rho(\Lambda)]\\ &\quad > x-2-(x-1)\exp[-1],\end{aligned}$$

which is positive for any $x > (2\exp[1]-1)/(\exp[1]-1) \approx 2.58$. However, in Footnote 11, we have shown that $x > 2.79$, which is sufficient for $a'(\Lambda) > 0$. In the same footnote, we show that $\lim_{\Lambda\to 1} x = 2\exp[\lambda_1]/\{\exp[\lambda_1]-1\}$, which in turn implies $a(1) = (x-2)/2 = 1/\{\exp[\lambda_1]-1\}$. In summary, we therefore conclude $X'(1) < a(1) = 1/\{\exp[\lambda_1]-1\}$, which is sufficient for $\lim_{\delta\to 1} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)} < 0$ to hold for any $\alpha \in (0,1)$. This implies that if the North chooses a tariff of $\tau = \overline{\tau}(1)$ it is a best response of the South to set $\delta < 1$.

Finally, since $\partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)}$ is a continuous function of δ and since $\lim_{\delta\to 0} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)} > 0$ and $\lim_{\delta\to 1} \partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta)} < 0$, the Intermediate Value Theorem guarantees existence of a $\delta^* \in (0,1)$ such that $\partial v_1/\partial\delta|_{\tau=\overline{\tau}(\delta^*)} = 0$. Thus, $\left(\delta^*, \overline{\tau}(\delta^*)\right)$ are mutual best policy responses of the two governments. This completes the proof.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00199-026-01740-1.

Acknowledgements We are grateful to the journal editor, two anonymous referees, and to Pao-Li Chang, Kerem Coşar, Harald Fadinger, Samuel Kortum, and Douglas Nelson for helpful comments and suggestions. We also thank participants at the annual conferences of the European Trade Study Group, the German Economic Association, and the Midwest Meetings; at the annual Göttingen Workshop in International Economics, the CESifo Global Economy Conference, and the 3rd TRISTAN Workshop; at the Trade & IO Workshop at the University of Bayreuth and the Workshop on Economic Security and the Future of the Global Trading System at the IfW Kiel; as well as seminar participants at CU Boulder and the University of Passau for valuable discussions. Hartmut Egger gratefully acknowledges financial support from the German

Research Foundation (DFG) through the Research Unit 'International Division of Labor: New Challenges for Workers, Firms, and Policy Makers" (Grant No. 550491480). Christian Fischer-Thöne acknowledges financial support from the Joachim Herz Foundation. A previous version of this manuscript circulated under the title "Trade Policy along the Global Value Chain: A Rationale for the Existence of Deep Trade Agreements."

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alfaro, L., Antràs, P., Chor, D., Conconi, P.: Internalizing global value chains: a firm-level analysis. J. Polit. Econ. **127**, 508–559 (2019)

Anderson, J.E., Larch, M., Yotov, Y.V.: Transitional growth and trade with frictions: a structural estimation framework. Econ. J. **130**, 1583–1607 (2020)

Antràs, P., Chor, D.: Organizing the global value chain. Econometrica **81**, 2127–2204 (2013)

Antràs, P., Chor, D.: Global Value Chains. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds)., Handbook of International Economics, Elsevier, vol. 5, chap. 5, 297–376 (2022)

Antràs, P., de Gortari, A.: On the geography of global value chains. Econometrica **88**, 1553–1598 (2020)

Antràs, P., Fort, T.C., Gutièrrez, A., Tintelnot, F.: Trade policy and global sourcing: a rationale for tariff escalation. J. Polit. Econ. Macroecon. **2**, 1–44 (2024)

Antràs, P., Staiger, R.W.: Offshoring and the role of trade agreements. Am. Econ. Rev. **102**, 3140–83 (2012)

Antràs, P., Staiger, R.W.: Trade agreements and the nature of price determination. Am. Econ. Rev. Pap. Proc. **102**, 470–476 (2012)

Bagwell, K., Bown, C.P., Staiger, R.W.: Is the WTO passé? J. Econ. Liter. **54**, 1125–1231 (2016)

Bagwell, K., Staiger, R.W.: Domestic policies, national sovereignty, and international economic institutions. Q. J. Econ. **116**, 519–562 (2001)

Bagwell, K., Staiger, R.W.: Can the Doha round be a development round? Setting a place at the table. In: Feenstra, R.C., Taylor, A.M. (eds.) Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century, pp. 91–124. University of Chicago Press, Chicago (2013)

Baldwin, R., Venables, A.J.: Spiders and snakes: offshoring and agglomeration in the global economy. J. Int. Econ. **90**, 245–254 (2013)

Belotti, F., Borin, A., Mancini, M.: icio: economic analysis with intercountry input-output tables. Stata J. **21**, 708–755 (2021)

Besedeš, T., Moreno-Cruz, J., Nitsch, V.: Depth and death: trade agreements and trade duration. Rev. Int. Econ. **33**, 519–536 (2025)

Beshkar, M., Lashkaripour, A.: The Cost of Dissolving the WTO: The Role of Global Value Chains, CAEPR Working Paper No. 2020-005 (2020)

Blanchard, E.J., Bown, C.P., Johnson, R.C.: Global value chains and trade policy. Rev. Econ. Stud. **93**, 181–214 (2026)

Bown, C.P., Sykes, A.O.: Trump's Reciprocal Tariffs Don't Add Up, Peterson Institute for International Economics (PIIE) Commentary, originally published in Bloomberg. (2017)

Caliendo, L., Feenstra, R.C., Romalis, J., Taylor, A.M.: A second-best argument for low optimal tariffs on intermediate inputs. J. Int. Econ. **145**, 103824 (2023)

Caliendo, L., Parro, F.: Trade Policy. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds)., Handbook of International Economics, Elsevier, vol. 5, chap. 4, 219–259 (2022)

Campolmi, A., Fadinger, H., Forlati, C.: Trade policy: home market effect versus terms-of-trade externality. J. Int. Econ. **93**, 92–107 (2014)
Campolmi, A., Fadinger, H., Forlati, C.: Trade and domestic policies under monopolistic competition. Econ. J. **135**, 892–913 (2025)
Conte, M., Cotterlaz, P., Mayer, T.: The CEPII Gravity Database, online available under http://www.cepii.fr/DATA_DOWNLOAD/gravity/doc/Gravity_documentation.pdf (2021)
Costinot, A., Vogel, J., Wang, S.: Global supply chains and wage inequality. Am. Econ. Rev. **102**, 396–401 (2012)
Costinot, A., Vogel, J., Wang, S.: An elementary theory of global supply chains. Rev. Econ. Stud. **80**, 109–144 (2013)
Egger, H., Kreickemeier, U.: Fairness, trade, and inequality. J. Int. Econ. **86**, 184–196 (2012)
Fliess, B., Arriola, C., Liapis, P.: Recent Developments in the Use of Export Restrictions in Raw Materials Trade, in Export Restrictions in Raw Materials Trade: Facts, Fallacies and Better Practices, ed. by OECD, OECD Publishing, 17–61 (2014)
Grossman, G.M., Helpman, E., Redding, S.J.: When tariffs disrupt global supply chains. Am. Econ. Rev. **114**, 988–1029 (2024)
Grossman, G.M., McCalman, P., Staiger, R.W.: The "new" economics of trade agreements: From trade liberalization to regulatory convergence? Econometrica **89**, 215–249 (2021)
Grossman, G.M., Rossi-Hansberg, E.: Trading tasks: a simple theory of offshoring. Am. Econ. Rev. **98**, 1978–1997 (2008)
Helpman, E., Itskhoki, O.: Labour market rigidities, trade and unemployment. Rev. Econ. Stud. **77**, 1100–1137 (2010)
Herkenhoff, P., Krautheim, S., Semrau, F.O., Steglich, F.: Corporate social responsibility along the global value chain. J. Dev. Econ. **167**, 103236 (2024)
Hofmann, C., Osnago, A., Ruta, M.: Horizontal Depth: A New Database on the Content of Preferential Trade Agreements, Policy Research Working Paper No. WPS 7981. (2017)
Hofmann, C., Osnago, A., Ruta, M.: The content of preferential trade agreements. World Trade Rev. **18**, 365–398 (2019)
Hsu, W.-T., Lu, L., Picard, P.M.: Income inequality, productivity, and international trade. Econ. Theor. **76**, 203–249 (2023)
Johnson, H.G.: Optimum tariffs and retaliation. Rev. Econ. Stud. **21**, 142–153 (1953)
Johnson, R.C.: Measuring global value chains. Annu. Rev. Econ. **10**, 207–236 (2018)
Johnson, R.C., Noguera, G.: A portrait of trade in value-added over four decades. Rev. Econ. Stat. **99**, 896–911 (2017)
Kennan, J., Riezman, R.: Do big countries win tariff wars? Int. Econ. Rev. **29**, 81–85 (1988)
Kohler, W.: International outsourcing and factor prices with multistage production. Econ. J. **114**, 166–185 (2004)
Kreickemeier, U., Qu, Z.: International trade with sequential production. Econ. Theor. **69**, 1101–1125 (2020)
Laget, E., Osnago, A., Rocha, N., Ruta, M.: Deep trade agreements and global value chains. Rev. Ind. Organ. **57**, 379–410 (2020)
Larch, M., Yotov, Y.V.: Estimating the effects of trade agreements: lessons from 60 years of methods and data. World Econ. **47**, 1771–1799 (2024)
Larch, M., Yotov, Y.V.: Deep trade agreements and FDI in partial and general equilibrium: a structural estimation framework. World Bank Econ. Rev. **39**, 281–307 (2025)
Lashkaripour, A., Lugovskyy, V.: Profits, scale economies, and the gains from trade and industrial policy. Am. Econ. Rev. **113**, 2759–2808 (2023)
Maggi, G., Ossa, R.: The political economy of international regulatory cooperation. Am. Econ. Rev. **113**, 2168–2200 (2023)
Mattoo, A., Mulabdic, A., Ruta, M.: Trade creation and trade diversion in deep agreements. Can. J. Econ. **55**, 1598–1637 (2022)
Melitz, M.J., Redding, S.J.: Missing gains from trade. Am. Econ. Rev. **104**, 317–321 (2014)
Noël, A.L.: How do sectoral compositions affect the impact of FTAs on global tariff liberalization? Econ. Theory, (2026). https://doi.org/10.1007/s00199-026-01718-z
OECD: OECD Inventory of Export Restrictions on Industrial Raw Materials 2025: Monitoring the Use of Export Restrictions Amid Growing Market and Policy Tensions, Tech. rep., OECD Publishing, Paris. (2025)

Orefice, G., Rocha, N.: Deep integration and production networks: an empirical analysis. World Econ. **37**, 106–136 (2014)
Osnago, A., Rocha, N., Ruta, M.: Do deep trade agreements boost vertical FDI? World Bank Econ. Rev. **30**, S119–S125 (2017)
Osnago, A., Rocha, N., Ruta, M.: Deep trade agreements and vertical FDI: the devil is in the details. Can. J. Econ. **52**, 1558–1599 (2019)
Ossa, R.: A new trade theory of GATT/WTO negotiations. J. Polit. Econ. **119**, 122–152 (2011)
Saggi, K., Wong, W.F., Yildiz, H.M.: Should the WTO require free trade agreements to eliminate internal tariffs? J. Int. Econ. **118**, 316–330 (2019)
Saggi, K., Wong, W.F., Yildiz, H.M.: The role of non-discrimination in a world of discriminatory preferential trade agreements. Can. J. Econ. **55**, 174–212 (2022)
Syropoulos, C.: Optimum tariffs and retaliation revisited: how country size matters. Rev. Econ. Stud. **69**, 707–727 (2002)
Wu, M.: Export Restrictions. In: Mattoo, A., Rocha, N., Michele, R. (Eds)., Handbook of Deep Trade Agreements, The World Bank, chap. 3, 89–110. (2020)
Yi, K.-M.: Can vertical specialization explain the growth of world trade? J. Polit. Econ. **111**, 52–102 (2003)
Zi, Y.: Trade costs, global value chains and economic development. J. Econ. Geogr. **20**, 249–291 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.