Preferential Trade Agreements and Global Sourcing*

Emanuel Ornelas John L. Turner Grant Bickwit

September 23, 2020

Abstract

We develop a new framework to study the welfare consequences of preferential trade agreements (PTAs) under global sourcing, incomplete contracts and endogenous matching. We uncover several new channels through which PTAs affect global welfare. Some effects stem from intensive margin changes—i.e., changes in investment and production in existing vertical chains—and from extensive margin relocations—i.e., due to the formation and destruction of vertical chains. In each case, there are potential trade-creating, trade-diverting and relationship-strengthening forces. The first two are reminiscent of the classical Vinerian approach, but take different forms under global sourcing. The third is entirely new in the regionalism literature and arises because PTAs affect the severity of hold-up problems in sourcing relationships. We characterize those forces and show circumstances when PTAs are necessarily welfare-enhancing or welfare-decreasing. In particular, we show that, because of the relationship-strengthening effect, PTAs can improve global welfare even when all types of trade-creation forces are absent.

Keywords: Regionalism; hold-up problem; sourcing; trade diversion; matching; incomplete contracts.

JEL classification: F13, F15, D23, D83, L22

---

*Ornelas: Sao Paulo School of Economics-FGV, CEPR, CESifo and CEP-LSE; eaornelas@gmail.com; Turner: Department of Economics, University of Georgia; jlturner@uga.edu. Bickwit: University of Georgia. We thank Emily Blanchard, Nuno Limão, Marcos Nakaguma, Michele Ruta, Bob Staiger, two anonymous referees, seminar participants at Australia National University, Banco Central de Mexico, Dartmouth, FED's Board, Hitotsubashi, INSEAD, LSE, Nottingham, Princeton, Singapore Management University, University of Tokyo, University of Hong Kong, several Brazilian universities (FGV-SP, PIMES, UFG, USP-RP), as well as conference participants at Stanford SITE 2016, InsTED 2017, LACEA 2017, REAP 2017, CEPR-ERWIT 2018, CES North America 2018, Fall 2018 Midwest International Trade, RIDGE 2018, TIGN 2018, ATW 2019, the 14th Arnoldshain Seminar on Institutions, Trade and Economic Policy, the Workshop on the Economics of International Trade Agreements at Villars-sur-Ollon, and the World Bank-CEPR Conference on Global Value Chains, Trade and Development for helpful comments and suggestions. We are especially grateful to Luis Araujo, Kala Krishna and Nathan Yoder for detailed discussions and suggestions. Ornelas acknowledges financial support from the Brazilian Council of Science and Technology (CNPq). Turner acknowledges financial support from the Terry-Sanford Research Program and the James C. Bonbright Center for the Study of Regulation.

1 Introduction

The past few decades have witnessed a sharp increase in the number of Preferential Trade Agreements (PTAs). Currently, the 164 members of the World Trade Organization have on average almost twenty PTA partners, whereas that figure was just over one in 1970.[1] A parallel trend has been the growth of trade in customized intermediate inputs and in the international fragmentation of production. As Johnson and Noguera (2017) document, the ratio of trade in value added to trade in gross exports has declined steadily in the 40 years up to 2009 for the manufacturing sector. Interestingly, they show that the decline was strongly influenced by reductions of bilateral trade frictions within PTAs. Indeed, Baldwin (2011, 2016), Blanchard (2015), Ruta (2017) and the World Trade Organization (2011), among others, have argued forcefully that global value chains (GVCs) are in reality mostly regional, driven by the formation of PTAs.

Strikingly, we lack even a basic framework to assess the desirability of PTAs in facilitating specialized input trade. This is what we provide in this paper. We consider a market with endogenous formation of two-firm vertical chains and non-contractible investments specific to relationships within chains. We show that the welfare analysis of regional integration in this context brings up several forces absent in traditional approaches. Welfare changes stem from the relocation of vertical chains across countries—an extensive-margin effect—and from different investment and production decisions in existing vertical chains—an intensive-margin effect.

Moreover, the investment and production incentives in some vertical chains that are formed are different from the investment and production incentives in some chains that are destroyed, so there are simultaneous changes in the two margins. This potentially complicates the welfare analysis significantly. Our first contribution is precisely to show how to strip the within-chain investment and production effects from the relocation effects. The resulting welfare decomposition provides a guideline to understand the welfare impact of PTAs in the context of global sourcing, and hence the conditions under which they are more likely to be beneficial in such settings.

Thus, keeping investment and production behavior in each chain unchanged, we show that the relocation effects display a parallel with Vinerian concepts of trade creation and trade diversion, even though they arise from a very different mechanism.[2] On the one hand, there is "matching creation," as vertical chains with domestic suppliers are replaced by others with more efficient suppliers in the PTA partner country. This raises global welfare. On the other hand, there is "matching diversion," as vertical chains with suppliers outside the bloc are replaced by others with inefficient suppliers in the PTA partner country. This lowers global welfare.

[1] For that calculation, we use the dataset constructed by Scott Baier and Jeffrey Bergstrand, available at https://www3.nd.edu/~jbergstr/ and first used by Baier et al. (2014).

[2] Since Viner (1950), analyses of preferential liberalization have typically relied on those concepts. Trade creation occurs when firms from partner countries produce more due to the PTA, at the expense of inefficient domestic firms. Trade diversion occurs when partner-country firms produce more due to the PTA, but at the expense of efficient nonmember firms. Those effects are based upon classical trade models, which rely on market clearing for price formation. That is also the approach taken in modern quantitative analyses of the welfare implications of PTAs (e.g., Caliendo and Parro, 2015), which are based on comparative-advantage forces, with anonymous markets and well-defined world prices for all goods. As will become clear, we depart from that approach in several dimensions.

Conversely, keeping the geographical structure of vertical chains fixed, we show that within-chain changes in investment and production behavior brought about by the agreement also generate two conflicting welfare effects. On the one hand, the tariff preference under a PTA induces the production of too many inputs within each vertical chain inside the bloc. This "sourcing diversion" lowers global welfare. On the other hand, the tariff preference boosts investment, which otherwise would be too low. This "relationship-strengthening" effect can raise global welfare, although it could also yield too much investment from a social standpoint. The total welfare consequence of a PTA is obtained when we combine all those effects.

Methodologically, we depart from the competitive environment to capture more accurately the realities of modern trade in intermediates. As Antràs (2020) stresses, they are very different, with the latter often involving specialized components that commit a buyer and a seller to each other. First, they need to find each other. Once matched, they become locked in the relationship and may underinvest in component-specific technology due to "hold-up problems" when contracts are incomplete (e.g., as in Grossman and Hart, 1986).

We introduce a property-rights model coupled with a Walrasian matching process to capture those effects in a tractable framework. The property-rights part of the model has been used before in related analyses (see discussion below). In contrast, our matching model has not been used in similar contexts before. It turns out to be particularly useful, because it permits rich comparative static analysis (easily represented in simple diagrams) on the effects of tariff level and country size on the number and productivity of active suppliers. Specifically, we have suppliers in different countries and with different levels of productivity matching with buyers to form vertical chains. Each supplier specializes her inputs to the buyer within their chain, and they bargain over terms of trade. Each buyer may source specialized inputs from within his chain and/or generic inputs from a competitive global market. The PTA affects matching, specialization investments and the composition of sourced inputs.

In the model, some domestic buyers form chains with suppliers from the partner country regardless of whether there is a PTA, while other suppliers there form chains with domestic buyers only when the PTA is in force. For the latter group, the key to understanding the implied welfare impact is the location of the previous matches of the domestic buyers. A match with a supplier in a non-PTA country yields a lower joint profit than a match with a supplier with the same productivity in a PTA country. Thus, when the bloc is formed, some buyers break vertical chains with existing suppliers outside the trading bloc to form chains with insiders. Because suppliers replaced outside the PTA are fundamentally more productive than those gained inside the PTA, welfare falls through that channel. That is what we call *matching diversion*—i.e., the PTA diverts matches away from productive suppliers outside the bloc to relatively unproductive suppliers inside the bloc.

At the same time, the PTA erodes a domestic supplier's advantage vis-à-vis imports from partner-country suppliers. Without the PTA, a match with a supplier abroad yields a lower joint profit than a match with a domestic supplier with the same productivity, but that is no longer true under a PTA. When the bloc is formed, some buyers then break vertical chains with existing domestic suppliers to form chains in the partner country. Because domestic suppliers replaced are fundamentally less productive than those gained inside the PTA, welfare rises through that channel. That is what we call *matching creation*—i.e., the PTA creates matches with more productive suppliers inside the bloc at the expense of relatively unproductive domestic suppliers.

Incentives also change for the vertical chains that include suppliers from the partner country both before and after the bloc. For those *incumbent suppliers*, the responses to preferential access generate a positive welfare effect provided that the severity of the hold-up problem is not extreme—if it is too severe, the investment incentive is too weak; if it is too mild, the investment incentive is excessive—and that the external tariff is not too high. Moreover, the welfare effect is higher whenever the distribution of supplier productivity is better, in the sense of stochastic dominance.

To understand the mechanisms, note that, under a PTA, incumbent suppliers receive a higher surplus on every unit traded. This propels more trade in specialized inputs, which in turn induces suppliers to increase their relationship-specific investments. Because without the PTA there is underinvestment due to a hold-up problem, the PTA-induced investment tends to improve efficiency. This relationship-strengthening effect always enhances welfare when the external tariff is sufficiently low, but causes excessive investment (and harms welfare) if the external tariff is too high. On the other hand, tariff discrimination introduces the usual welfare-reducing diversion effect—here, too many expensive specialized inputs are sourced—which increases monotonically in the tariff. Importantly, this sourcing diversion is independent of the number of units the firms in a vertical chain initially trade with each other. In contrast, since the investment yields greater value to every unit traded, the relationship-strengthening effect is stronger, the more units the firms initially trade. Therefore, it is more likely to dominate the negative sourcing-diversion effect when firms trade large volumes of specialized inputs even without the PTA—i.e., when they have high productivity.

It is important to note that, while we have the incomplete-contract analog of trade diversion at both the intensive and the extensive margins, and the incomplete-contract analog of trade-creation at the extensive margin, we design the model to shut down trade creation forces at the intensive margin. We put it aside to shed light on potentially important forces so far ignored in the academic literature and in policy circles alike, here encapsulated in what we define as the relationship-strengthening effect. By doing so we make it clear that any positive within-match welfare effect stems from the novel relationship-strengthening effect.[3]

Having characterized all forces at play, we look at circumstances when positive or negative effects would unambiguously dominate. For example, we show that when the external tariff is sufficiently low, a PTA in necessarily welfare-enhancing. The reason is that matching creation, matching diversion and sourcing diversion become second-order small, but the relationship-strengthening effect is first-order positive for low external tariffs. Under some parameter restrictions, we show that the welfare effect of a PTA is an inverted-U function of the external tariff and is positive if

[3] In section 7.1 we show how the model would need to be altered and how results would change if we allowed for trade creation at the intensive margin.

and only if the external tariff is sufficiently low.

We also look at specific types of PTAs. For the case of "natural trading partners," in which the agreement is between countries that trade heavily with each other, a PTA always enhances welfare provided that the intensive-margin effect is positive. The reason is that, in this case, the extensive-margin effect consists of only matching creation. This result provides a possible rationale for the "natural trading partners hypothesis," which posits that those types of agreements are more likely to enhance welfare.[4]

We also consider "North-South" agreements, in which countries specialize in different stages of the production processes. This case stacks the odds against making PTAs welfare-enhancing. The home country has no suppliers, so there is no matching-creation effect. Indeed, there are no trade-creation effects at all, so welfare can rise only if relationship-strengthening effects dominate. For sufficiently high tariffs, relationship-strengthening effects are dominated in both the intensive and extensive margins, and PTAs necessarily reduce welfare. For sufficiently low tariffs, however, the intensive-margin effect is positive. Perhaps most surprisingly, even the extensive-margin effect alone *may* be positive, provided that relationship-strengthening effects dominate both sourcing-diversion and matching-diversion effects.

Overall, our paper illustrates how global sourcing can fundamentally change the normative implications of PTAs, sometimes entirely reversing Viner's (1950) original idea: even purely trade-diverting PTAs can be helpful when one considers how they can mitigate hold-up problems created by incomplete contracts. The central point is that, when it comes to the trade of customized inputs, tariff preferences do not affect just allocative efficiency; they also affect the efficiency of the production process, through changes in the incentives to invest and to form vertical chains. The upshot is that the welfare implications of PTAs under global sourcing are much more subtle and intricate than standard models suggest.

Now, while our model contrasts with standard regionalism theories in its motivation, its mechanisms and its results, it relates to several branches of the trade literature. In terms of structure, we build on Ornelas and Turner (2008, 2012), where we develop analyses under incomplete contracts and relationship-specific investments. Our goal in those papers is, respectively, to explain the worldwide growth of trade flows relative to GDP until 2007 and to characterize optimal unilateral trade policies when firms' organizational mode is endogenous, whereas our current focus on preferential liberalization is entirely new. More crucially, the previous papers carried out the analysis from the perspective of a single vertical relationship, and did not consider heterogeneity in productivity or endogenous matching. Those are essential ingredients of the current analysis—without them, neither the intensive-margin nor the extensive-margin effects uncovered here could even be analyzed.

In terms of economic environment, the paper is closely related to Antràs and Staiger (2012a, b) and to a recent paper by Grossman and Helpman (2020). The goal of Antràs and Staiger (2012a, b)

[4] The natural trading partners hypothesis is often relied upon in policy circles and has empirical support (e.g., Baier and Bergstrand, 2004), but lacks solid theoretical foundations (see Bhagwati and Panagariya, 1996).

is to study the optimal design of (nondiscriminatory) trade agreements, not an issue we address, but their more general point is to show that the efficiency properties of international trade agreements are vastly different when buyers/consumers and sellers/producers must negotiate their terms of trade through bargaining. That may be a consequence of hold-up problems and/or matching, but the key element in their analyses is the absence of market-clearing conditions fully disciplining world prices. That is also a central element in our analysis. Our model structure is, however, very different from Antràs and Staiger's (2012a, b); hence, we develop different mechanisms and generate entirely novel results. In particular, we underscore how tariff preferences shape the structure of production through their effects on investment and matching decisions.

In turn, Grossman and Helpman (2020) study how tariffs affect the structure of GVCs in a setting where each buyer searches for, matches and bargains with suppliers of different components. They provide a rich analysis of how tariffs can affect the incentives for searching and the outcome of bargaining in those relationships, which they use to assess the welfare implications of tariffs. The net effect is ambiguous due to multiple offsetting forces, but they show that, under plausible parameter values, the importing country is likely to lose with a tariff. Such losses become particularly prominent once the (discriminatory) tariff becomes large enough to induce partial relocation of supply chains to high-cost countries. That effect is akin to what we deem matching diversion in our setting. The key difference is that, in Grossman and Helpman (2020), this extensive-margin effect is necessarily negative, whereas here it is coupled with offsetting incentives to invest and produce.

Our paper also complements research using detailed models of intermediate input trade and bargaining in international trade.[5] In particular, it shares important characteristics with the analysis of Grossman and Helpman (2005), which also features a choice of location for outsourcing decisions as well as matching with suitable suppliers. The structures of the models are quite different, however. For example, whereas Grossman and Helpman adopt an "all-or-nothing" specification for the relationship-specific investments, in our setup investments are continuous, implying that in the absence of trade agreements investment is always suboptimal. More importantly, the goals of the analyses are completely distinct. For instance, the role of market thickness in shaping outsourcing decisions features prominently in Grossman and Helpman (2005), whereas we concentrate on the themes described above.

Our study adds as well to the literature that links trade liberalization to investment and innovation. That line of research is best exemplified by Bustos (2011) and Lileeva and Trefler (2010), who provide compelling theoretical analyses combined with empirical support for their models' predictions. In both papers, the empirical analysis relies on the reduction of preferential tariffs (Argentine firms facing lower tariffs in Brazil under Mercosur in one case, Canadian firms facing lower tariffs in the U.S. under CUSTA in the other), although their models pay no heed to the preferential nature of the liberalization. In contrast, we stress precisely the discriminatory aspect

[5] This line of research includes, among others, McLaren (2000), Qiu and Spancer (2002), Antràs and Helpman (2004, 2008), Antràs and Chor (2013) and Antràs, Fort and Tintelnot (2017).

of tariff changes, and we analyze how they affect investment and matching patterns in international sourcing decisions, not a special concern of Bustos (2011) and Lileeva and Trefler (2010).[6]

Finally, the paper contributes to a large literature on regional trade agreements, in particular the strand that focuses on the welfare implications of preferential integration. For recent surveys, see Bagwell, Bown and Staiger (2016), Freund and Ornelas (2010), Limao (2016) and Maggi (2014).

The paper is organized as follows. We set up the basic model in section 2. We then study the equilibrium within a given vertical chain in section 3 and characterize the matching equilibrium in section 4. We assess each of the forces behind the welfare consequences of a PTA in section 5, while in section 6 we analyze its overall welfare effects, in general and in some particular cases. In section 7 we discuss extensions and testable implications of our model. We conclude in section 8.

2 Model

There is a continuum of differentiated final goods available for consumption in the world economy. Consumption of those goods increases the utility of consumers at a decreasing rate, and there are no cross-product effects. There is also an homogeneous numéraire good y that enters consumers' utility function linearly. Thus, if they purchase any amount of y, any extra income will be directed to the consumption of the numéraire good. We assume relative prices are such that consumers always purchase some good y. Furthermore, production of one unit of y requires one unit of labor, the market for good y is perfectly competitive, and y is traded freely. This sets the wage rate in the economy to unity.

All the action happens in the differentiated sector. For each differentiated final good, production requires transforming intermediate inputs under conditions of decreasing returns to scale. Production is carried out by buyer (B) firms located in the *Home* country. Those firms act as aggregators, transforming intermediate inputs, all produced only with labor, into marketable goods. Final good producers obtain net revenue $V(Q)$ when they buy and process Q intermediate inputs, where $V' > 0$ and $V'' < 0$. Under this structure, there are no general equilibrium effects across sectors. Thus, without further loss of generality, we develop the analysis as if there were a single differentiated sector.

There is another country, *Foreign*, as well as the rest of the world (ROW). When sourcing, each buyer may purchase generic inputs z available in the world market and/or customized inputs

[6] In related research, Conconi, Garcia-Santana, Puccio and Venturini (2018) show empirically how NAFTA's rules of origin (ROOs) affected the pattern of sourcing within the bloc. Although we abstract from ROOs in our analysis, our framework could be adjusted to assess their welfare consequences, as we discuss in the conclusion. Also related is the paper by Blanchard, Bown and Johnson (2017). They analyze, theoretically and empirically, optimal trade policy in the context of GVCs, an issue we sidestep here, but which could be studied in a modified version of our framework. Heise, Pierce, Schaur and Schott (2017) study as well how trade policy alters international patterns of procurement, but their proposed mechanism—how changes in trade policy uncertainty affect the mode of sourcing relationships—is very different from ours. From a different angle, Antràs and de Gortari (2020) develop a general equilibrium framework to study how exogenous trade costs shape the geography of GVCs. Their focus is on characterizing how production and trade costs along the value chain shape the equilibrium structure of GVCs. PTAs are likely to be an important component of that cost structure, as Johnson and Noguera (2017) argue.

q from a specialized supplier (S). Generic inputs are produced by a competitive fringe and require p_w units of labor. Thus, their price in the world market is p_w. We consider that *Home* is too small to affect p_w. *Home*'s buyers face a per-unit tariff t on all imported intermediate goods, so a generic imported input costs $p_w + t$ for them. One unit of generic input and one of specialized input have the same revenue-generating value for a buyer.[7] Under this normalization, B's revenue depends only on the total number of intermediate inputs he purchases, Q, and not on its composition.[8]

We set the tariff on q and z to be the same, t, but this is for expositional simplicity only. It is straightforward to allow for different tariffs on q and z. We also assume that neither *Home* nor *Foreign* produces generic inputs. That is not without loss of generality, but helps us convey our main ideas in a simple way. In section 7.2 we discuss how alternative configurations of the generic industry would affect our analysis. As we explain there, specific results would be (obviously) different under different patterns of trade, but the thrust of the analysis would remain unchanged.

Now, to acquire specialized inputs, a buyer must first match with a supplier and form a vertical chain. There is a unit mass of heterogeneous suppliers in the world and a mass of size $\beta \in (0, 1)$ of identical buyers in *Home*.[9] Suppliers are split between *Home*, *Foreign* and *ROW* proportionally to γ_H, γ_F and $1 - \gamma_H - \gamma_F$, respectively. Each supplier is identified by ω, a heterogeneity parameter that indexes (the inverse of) her productivity. The distribution of suppliers in each country follows a continuous and strictly increasing distribution $G(\omega)$, with an associated density $g(\omega)$, where ω lies on $[0, p_w]$.[10] To focus on fundamental forces, we consider a simple Walrasian matching environment where each supplier who matches pays a fee to her buyer. We will see that, in that setting, the equilibrium matching structure follows efficient sorting—i.e., low-ω suppliers match but high-ω suppliers do not—and is stable. We further assume that $t < \min\{p_w - G^{-1}(\beta), G^{-1}(\beta)\}$: this ensures that, in all countries, the most-productive supplier always succeeds in matching with a buyer and the least-productive supplier never succeeds in matching with a buyer.

Upon forming a vertical chain, B and S specialize their technologies toward each other. This specialization costs nothing, but implies that at any point a buyer purchases specialized inputs from only one supplier. After B and S specialize toward each other, S pays for a non-contractible relationship-specific investment that lowers her marginal cost prior to trade with B. The investment is observed by both B and S, but is not verifiable in a court of law. Nothing essential would change if the buyer also made an analogous ex-ante investment.

[7] Note that this is just a normalization. To see that, suppose that buyers value generic and specialized inputs differently. Then add a multiplicative 'compatibility cost' to the use of generic inputs. Call such costs ξ. That would increase the quality-adjusted cost of generics for their buyers to $\xi p_w + t$. But we could then simply redefine units by dividing the units of generic inputs by ξ and adjusting the tariff accordingly.

[8] The assumption of perfect substitutability between q and z (adjusted for quality) simplifies the analysis significantly, but is not essential. However, it is critical that they are substitutes to some degree.

[9] Buyer heterogeneity is an important empirical phenomenon, which has been documented and analyzed by several authors (see, e.g., Antràs et al., 2017; Blaum et al., 2018; Bernard et al., 2018; Sugita et al., 2018). We abstract from that additional source of heterogeneity to keep the analysis tractable.

[10] As it will become clear shortly, in the absence of preferential treatment, specialized inputs are not provided when $\omega > p_w$, as in that case the buyer-supplier pair would gain nothing by trading. Since in equilibrium all suppliers with $\omega > p_w$ do not specialize, it is useful to limit the analysis to the more interesting case where the upper limit of the distribution of suppliers is p_w, and $G(\omega)$ is the truncated distribution of suppliers when $\omega \leq p_w$.

Once investment is sunk, the firms decide how much to trade and at what price. The specialized inputs are not traded on an open market, and have no value outside the chain. Furthermore, the parties cannot use contracts to affect their trading decisions.[11] Instead, they need to bargain over price and quantity of specialized inputs. If bargaining breaks down, S produces the numéraire good and earns zero (ex-post) profit, while B purchases only generic inputs. If bargaining is successful, B imports generic inputs from ROW and purchases specialized inputs from S. Finally, B transforms all inputs into a differentiated final good and payoffs are realized.

To generate clear-cut analytical solutions, we adopt some specific functional forms. Conditional on investment i, we specify the cost function for specialized units q as

$$C(q,i,\omega) = (\omega - bi)q + \frac{c}{2}q^2. \tag{1}$$

Parameter ω is the intercept of the marginal cost curve; the lower is ω, the more efficient is the supplier. The slope of marginal cost is c, while the effectiveness of investment in reducing production costs is b. In turn, the investment cost is

$$I(i) = i^2.$$

We assume that $2c > b^2$.[12]

Concrete functional forms are useful to analyze PTAs, where changes in tariffs are not marginal but discrete, from their initial levels to zero, and because we want to condition results on the extent of the margin of preference. Our linear-quadratic specification displays properties that are standard and provide a good representation of the key elements of our environment: investment and productivity reduce both cost and marginal cost ($C_i < 0, C_{qi} < 0, C_\omega > 0, C_{q\omega} > 0$); the marginal cost curve is positively sloped ($C_{qq} > 0$) but its slope can vary (c is a parameter); and the cost of investment is convex ($I' > 0, I'' > 0$). This specification has the advantage of permitting full analytical solutions at the level of a single buyer-supplier pair, a straightforward analysis of Walrasian matching with and without a PTA, and a precise welfare analysis.[13]

We focus on the case where B engages in dual sourcing, purchasing both generic and specialized inputs. Define Q^* as the equilibrium level of total inputs sourced. When B imports some generic inputs, his marginal gain from that purchase, $V'(Q^*)$, must equal his marginal cost, $p_w + t$; this pins down Q^*. To ensure production of the final good, the initial level of marginal revenue for B needs to be sufficiently high: $V'(0) > p_w + t$. To ensure that S does not produce all inputs,

[11] This would be the case, for example, if quality were not verifiable in a court and the supplier could produce either high-quality or low-quality specialized inputs, with low-quality inputs entailing a negligible production cost for the seller but being useless to the buyer. This is the same approach used by Antràs and Staiger (2012a), among others.

[12] This ensures that the sensitivity of marginal cost to investment is not too large. If b were too large, not satisfying the condition, every supplier would want to make $i \to \infty$.

[13] Naturally, the functional forms do impose restrictions. In particular, (1) implies $C_{qqq} = 0$, so the marginal cost curve has no curvature. While this is a very common assumption in international trade models (which often assume further that $C_{qq} = 0$), the sharpness of some of our results does depend on $C_{qqq} = 0$. Effectively, they follow through if C_{qqq} and I''' are sufficiently small in absolute value, but the analysis becomes particularly clean if one sets them to zero, as we do here.

we assume $C_q(V'^{-1}(p_w + t), b(p_w + t)\left(2c - b^2\right)^{-1}, 0) > p_w + t$. This simply implies that even the marginal cost of the most productive firm ($\omega = 0$) is high enough so that B wants to purchase some generic inputs. In addition to being realistic,[14] the dual sourcing specification is pedagogical, as will soon become clear. More generally, the important requisite is that the buyer must have the *option* of buying generics when negotiating with his specialized supplier, because that establishes the threat point in the bargaining process. Nevertheless, in section 7.1 we show how the analysis would change in cases where single-sourcing becomes optimal.

Figure 1 illustrates the Y-shaped supply chain in this economy. To distinguish from the B-S vertical chain, we use the term *Y-chain* when referring to the entire supply chain. Naturally, this is a very simple production chain, whereas in reality they can be (and often are) longer and more complex.

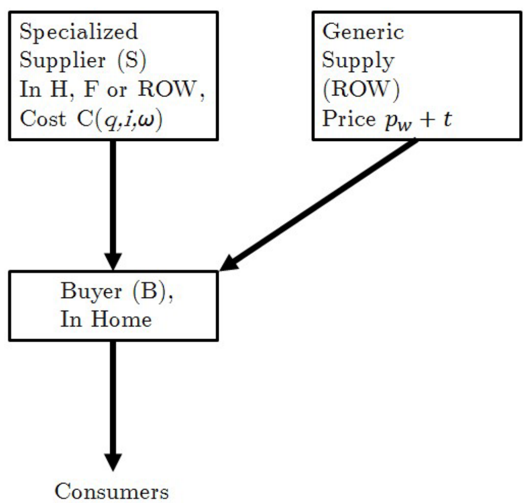

Fig. 1: Y-chain

Notes: The figure shows the relationship between a single buyer and its two sources of input supply.

The timing of events is as follows. First, (i) each B matches with a supplier S in one of the three countries to form a vertical chain, adapting their technologies toward each other within the chain. Then, (ii) S makes an irreversible relationship-specific investment. Once the investment is sunk, (iii) B and S bargain over price and quantity of q. If bargaining is successful, trade of q takes place and payments are made; otherwise, $q = 0$ and S produces the numéraire good. Subsequently, (iv) B purchases z. Finally, (v) production occurs and final goods are sold.

We solve the game by backward induction. First, we carry out the analysis from the perspective of a single vertical chain. We then solve for the equilibrium structure of matches.

[14] Mixing customized and standardized inputs is a rather common practice, as for example Boehm and Oberfield (2020) document for Indian manufacturing plants.

3 Single Vertical Chain

There are two possible trade regimes: no trade agreement and a PTA between *Home* and *Foreign*. When there is no trade agreement, all inputs imported into *Home* are subject to the tariff regardless of their origin, whereas specialized inputs produced in *Home* do not incur tariffs. Thus, without a PTA, t represents the tariff advantage that *Home* specialized suppliers have over suppliers in *Foreign* and ROW.

Under a PTA, the tariff on goods traded between *Home* and *Foreign* is eliminated. Imports from ROW still face tariff t, which is now the *external* tariff under the agreement, assumed unchanged. Thus, t also represents the preferential margin offered to imports coming from *Foreign* under the PTA, relative to those coming from ROW.

Hence, with respect to tariff treatment in the *Home* market, a supplier may or may not receive preferential treatment. *Home* specialized suppliers are always in the former group, whereas ROW suppliers are always in the latter. Tariff treatment varies across trade regimes only for *Foreign* specialized suppliers, who receive preferential treatment under a PTA but not otherwise.

3.1 Vertical Chains without Preferential Treatment

We first consider the vertical chains that do not receive preferential treatment. After S chooses her investment, B and S determine the number and price of the specialized intermediate inputs by Generalized Nash Bargaining. Specifically, let the supplier have bargaining power $\alpha \in (0,1)$. Under Generalized Nash Bargaining, the two firms choose the number of specialized inputs q and their price p^s to maximize

$$\left[U^S(Barg) - U^S(NoBarg)\right]^{\alpha} \left[U^B(Barg) - U^B(NoBarg)\right]^{(1-\alpha)},$$

where $U^j(m)$ is the gross profit (i.e., profit absent transfers) that firm j (either B or S) would receive under scenario m. The two possible scenarios are either bargaining and trading ($m =$*Barg*) or not reaching a bargain and thus not trading ($m =$*NoBarg*).

We use subscript 0 to indicate variables in Y-chains where the B-S vertical chain is unprotected by the tariff. Conditional on inverse productivity ω, investment i, specialized inputs q and generic inputs z (with $q + z = Q^*$), utilities are as follows: $U_0^B(Barg) = V(Q^*) - (p_w + t)z - (p^s + t)q$; $U_0^B(NoBarg) = V(Q^*) - (p_w + t)Q^*$; $U_0^S(Barg) = p^s q - C(q, i, \omega) - I(i)$; and $U_0^S(NoBarg) = -I(i)$.

Define $\Sigma_0 \equiv \left[U_0^S(Barg) - U_0^S(NoBarg)\right] + \left[U_0^B(Barg) - U_0^B(NoBarg)\right]$ as the bargaining surplus with no preferential treatment. The outcome of bargaining has the two firms splitting the proceeds, with S receiving $\alpha\Sigma_0$ and B receiving $(1-\alpha)\Sigma_0$, in addition to their reservation payoffs. Conditional on i and q, this bargaining surplus is

$$\Sigma_0 = p_w q - C(q, i, \omega). \tag{2}$$

Conditional on i and t, efficient bargaining dictates that a B-S vertical chain trade the ex-post privately optimal amount of specialized inputs, defined as q_0, and B purchases the ex-post privately efficient level of generic inputs, defined as z_0. Together, they determine the total number of inputs purchased within the Y-chain by B: $Q^* = q_0 + z_0$. Because both specialized and generic inputs incur the tariff, privately optimal sourcing equalizes the marginal cost of the two alternative inputs:

$$p_w = C_q(q_0, i, \omega). \tag{3}$$

This pins down q_0 (and hence z_0) for given i. Under our cost specification, this condition is

$$q_0 = \frac{p_w - \omega + bi}{c}. \tag{4}$$

Now, anticipating the bargaining outcome, S chooses her investment by solving

$$\max_i \alpha\Sigma_0 - I(i).$$

Thus, equilibrium investment, i_0^*, satisfies $I'(i_0^*) = -\alpha C_i(\cdot)$, or equivalently,

$$i_0^* = \left(\frac{\alpha b}{2c - \alpha b^2}\right)(p_w - \omega). \tag{5}$$

Substituting (5) back in (4) and manipulating, we find

$$\begin{aligned} q_0^* &= \left(\frac{2}{\alpha b}\right)\left(\frac{\alpha b}{2c - \alpha b^2}\right)(p_w - \omega) \\ &= \left(\frac{2}{\alpha b}\right) i_0^*. \end{aligned} \tag{6}$$

Hence, the equilibrium investment and output are proportional. More productive (lower-ω) suppliers produce more for a given investment, and they also invest more, reinforcing their original cost advantage. When supplier bargaining power (α) is low, equilibrium investment is low, and it drops to zero as $\alpha \to 0$, when S does not appropriate any surplus created by the investment. As α increases, both investment and production of specialized inputs increase. They are also positively affected by the effectiveness of investment (b), but negatively affected by the steepness of the marginal cost curve (c). Observe that neither investment nor production is affected by the tariff, which distorts the total volume of inputs, Q^*, but does not interfere with the sourcing of q.

It is useful to compare i_0^* with the efficient level of investment. Under privately efficient sourcing, worldwide social welfare due to this bilateral relationship can be defined as

$$\Psi_0 = V(Q^*) - p_w Q^* + p_w q_0^* - C(q_0^*, i_0^*, \omega) - I(i_0^*). \tag{7}$$

Observe that Ψ_0 includes three components: reservation payoffs $[V(Q^*) - (p_w + t)Q^*$ for B, $-I(i_0^*)$ for $S]$; bargaining surplus $[p_w q_0^* - C(q_0^*, i_0^*, \omega)]$; and tariff revenue $[tQ^*]$.

The *efficient* level of investment (i^e) maximizes (7). Under dual sourcing, the first two terms of (7) are unaffected by the level of investment. Thus, using (3), it follows that efficiency requires

$$I'(i^e) = -C_i(\cdot). \tag{8}$$

Under our cost specifications, this yields

$$i^e = \left(\frac{b}{2c - b^2}\right)(p_w - \omega). \tag{9}$$

Comparing i_0^* with i^e, it is immediate that $i_0^* < i^e$ (since $\alpha < 1$). Moreover, it is easy to see that the extent of the hold-up problem, which we define as

$$HUP_0 \equiv i^e - i_0^*, \tag{10}$$

depends critically on the productivity of the supplier (all proofs not shown in the main text are in the Appendix):

Lemma 1 *The extent of the hold-up problem in the absence of preferential treatment, HUP_0, increases with S's productivity (i.e., as ω falls).*

Intuitively, the equilibrium level of investment increases with S's share α of the bargaining surplus, whereas the efficient investment increases with the whole bargaining surplus. The inefficiency is therefore proportional to $(1-\alpha)\Sigma_0$, which increases with S's productivity. Hence, the vertical chains with the best suppliers—who produce more and generate higher surplus for any level of investment—are more negatively affected by contract incompleteness. Observe also that $HUP_0 = 0$ when $b = 0$ and rises with b, as b increases the effectiveness of investment.

We can solve for closed-form expressions for equilibrium gross profits conditional on ω:

$$U_0^S(\omega) = \frac{\alpha (p_w - \omega)^2}{2c - \alpha b^2}, \tag{11}$$

$$U_0^B(\omega) = [V(Q^*) - (p_w + t)Q^*] + \frac{2c(1-\alpha)(p_w - \omega)^2}{(2c - \alpha b^2)^2}. \tag{12}$$

Both are clearly decreasing in ω, so low-ω suppliers earn higher gross profits than high-ω suppliers, and a buyer's gross profit is higher in a vertical chain with a low-ω supplier. These values are realized for B-S vertical chains where the supplier is either in *ROW* or in *Foreign* under no PTA.

3.2 Vertical Chains with Preferential Treatment

We now consider the vertical chains that receive preferential treatment, whose specialized supplier is either in *Home* or in *Foreign* under a PTA. The total volume of inputs purchased by B remains unchanged at Q^*, as pinned down by $V'(Q^*) = p_w + t$, but now its composition reflects different relative prices of specialized and generic inputs. Under bargaining, just one of the potential $U^j(m)$

payoff terms, $U^B(Barg)$, structurally changes, becoming (for price p^s)

$$U_t^B(Barg) = V(Q^*) - (p_w + t)z - p^s q,$$

where we use subscript t to indicate variables in Y-chains where the B-S vertical chain is protected by the tariff. The bargaining surplus under preferential treatment, Σ_t, also reflects the change in buyer gross profit with trade due to tariff savings when B sources from S. Conditional on q and i, we have

$$\Sigma_t = (p_w + t)q - C(q, i, \omega).$$

Conditional on i and t, efficient bargaining dictates that a B-S vertical chain trade the ex-post privately optimal number of specialized inputs, defined as q_t, and B purchase the ex-post privately efficient level of generic inputs, defined as z_t. Because specialized inputs do not incur the tariff but generic inputs do, privately optimal sourcing satisfies the following condition:

$$p_w + t = C_q(q_t, i, \omega). \tag{13}$$

This pins down q_t (and hence z_t) for given i. Under our cost specification, this yields

$$q_t = \frac{p_w + t - \omega + bi}{c}. \tag{14}$$

The investment decision solves

$$\max_i \alpha \Sigma_t - I(i),$$

so equilibrium investment is

$$i_t^* = \left(\frac{\alpha b}{2c - \alpha b^2}\right)(p_w + t - \omega). \tag{15}$$

We define the change in investment due to preferential treatment as $\Delta i \equiv i_t^* - i_0^*$. Our quadratic cost specification yields the useful property that Δi is proportional to the tariff:

$$\Delta i = \left(\frac{\alpha b}{2c - \alpha b^2}\right) t. \tag{16}$$

The change in investment vanishes when $\alpha \to 0$ and is strictly increasing in α. Similarly, it increases with the responsiveness of marginal cost to investment (b) and is zero if marginal cost is completely unresponsive ($b = 0$). Finally, Δi decreases with the slope of the marginal cost curve (c).

The resulting equilibrium level of specialized inputs remains proportional to investment,

$$q_t^* = \left(\frac{2}{\alpha b}\right) i_t^*, \tag{17}$$

and the change in specialized inputs, $\Delta q \equiv q_t^* - q_0^*$, is also proportional to Δi:

$$\begin{aligned}\Delta q &= \left(\frac{2}{2c - \alpha b^2}\right) t \\ &= \left(\frac{2}{\alpha b}\right) \Delta i.\end{aligned}$$

Part of the increase in the quantity, $\frac{t}{c}$, is due entirely to S's discriminatory protection, and occurs even with no additional investment. To see that, notice that when $b = 0$ the supplier never invests and yet sales of specialized inputs still increase, by exactly $\Delta q(b = 0) = \frac{t}{c} > 0$.

The sales of specialized inputs increase also because of lower production costs. When investment is higher, S's entire marginal cost curve is lower. From an efficiency standpoint, S *should* produce level q_1^*, satisfying $C_q(q_1^*, i_t^*, \omega) = p_w$. Developing this expression under our cost specification and using (4), we obtain

$$\begin{aligned}q_1^* &= q_0^* + \left(\frac{\alpha b^2}{2c - \alpha b^2}\right) \frac{t}{c} \\ &= q_0^* + \frac{b}{c} \Delta i.\end{aligned}$$

It is easy to see that

$$q_t^* = q_1^* + \frac{t}{c}.$$

That is, S produces $\frac{t}{c}$ units more than it should, from an efficiency standpoint.

Figure 2 highlights the effects, within a single Y-chain whose supplier is in *Foreign*, of granting tariff preferences under a PTA. Units $q \in (0, q_0^*]$ are sold regardless of whether there is a PTA. But due to the higher investment, there is extra bargaining surplus for each of those units, because S's marginal cost is lower. This extra surplus is shown by area **C**. Units $q \in (q_0^*, q_1^*]$ are produced by S under the PTA, but not otherwise. They represent trade driven by productivity growth. The additional surplus from those units is shown by area **D**. The $\frac{t}{c}$ units produced by S under the PTA at a marginal cost higher than p_w are those between q_1^* and q_t^*. They reflect classic trade diversion, as the extra specialized inputs come at the expense of generic inputs. That extra production leads to the deadweight loss shown by area **E**. Furthermore, under a PTA there is also an additional investment cost (not shown in the figure), which reduces the overall welfare gain.[15]

Interestingly, the PTA can lead to *too much* investment relative to the efficient level. Recall that, without the agreement, $HUP_0 = i^e - i_0^* > 0$. Such an unambiguous ordering between equilibrium and efficient levels of investment does not exist under preferential treatment. Defining the excess

[15] Observe that a change in parameter ω provokes a parallel shift of the marginal cost curve. It is easy to see that such a shift does not affect the size of area **E**, which is therefore independent of the supplier's productivity. Similarly, because the change in investment is also unaffected by ω (see equation 16), a lower ω causes the same parallel shift of the two C_q curves in Figure 2. As a result, the size of area **D** is also independent of the supplier's productivity. On the other hand, area **C** is decreasing in ω, since a lower ω increases q_0^*.

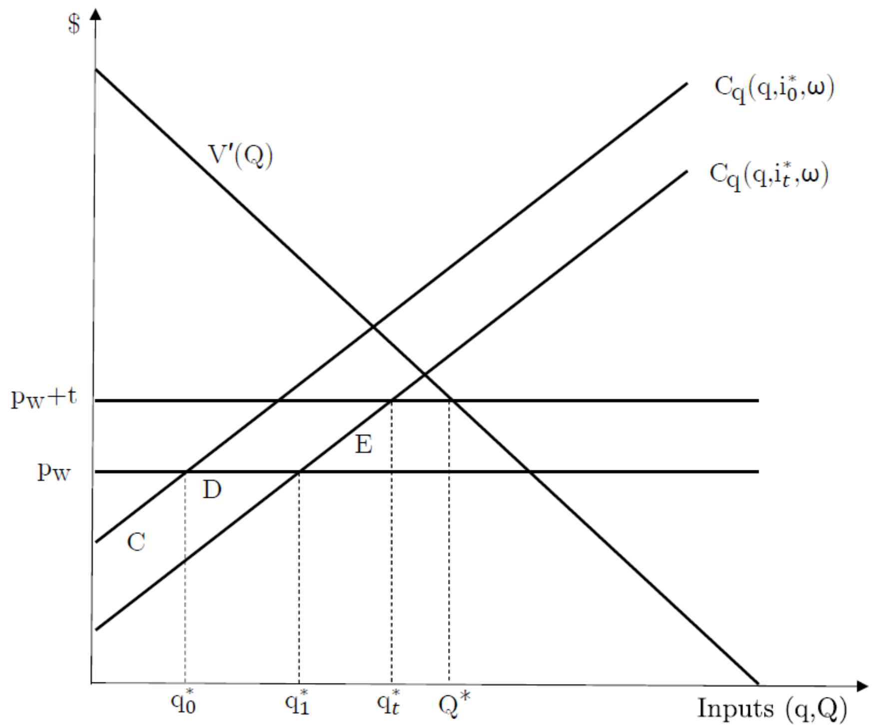

Fig. 2: The Effects of a PTA on Sourcing and Production within a Y-chain

Notes: The figure shows (some of) the welfare effects of a PTA within a Y-chain. Area C represents gains due to lower production costs for units already traded. Area D represents gains due to lower production costs for new units traded within the chain. Area E represents losses due to trade diversion in inputs. The figure omits changes in investment costs, which are higher under the PTA.

of investment under preferential treatment as

$$EXC_t \equiv i_t^* - i^e, \tag{18}$$

one finds that

$$EXC_t > 0 \Longleftrightarrow (2c - b^2)\alpha t > 2c(1-\alpha)(p_w - \omega).$$

It follows that $i_t^* > i^e$ when α is sufficiently close to unity (in which case the original hold-up problem is relatively unimportant, so Δi is mostly distortionary) and/or when t is sufficiently high (in which case Δi is too high).[16] Also, excess investment is lower for higher-productivity suppliers.

Lemma 2 *The extent of excess investment with preferential treatment, EXC_t, falls with S's productivity (i.e., as ω falls).*

Overall, this analysis highlights a "within Y-chain" trade-off between conventional trade/sourcing

[16] In the Appendix we show that the efficient level of investment is the same with and without a PTA.

diversion and an effect that so far has been entirely neglected in the regionalism literature. Due to the PTA, a chain in *Foreign* creates additional surplus for all units of specialized inputs that would be produced without the agreement—area **C** in Figure 2—plus some surplus for additional units traded—area **D**. This may increase welfare, if it more than offsets the losses due to excessive production (area **E**) and to additional investment.

It is important to stress at this point that, while our model displays an effect akin to Vinerian trade diversion within Y-chains, it shuts down Vinerian trade creation within Y-chains by keeping total units traded, Q^*, fixed (for given t). We specify the model this way both for analytical convenience and to shed light on novel channels through which PTAs affect economic efficiency. Nevertheless, in section 7.1 we discuss how our analysis would change if we allowed for within-chain trade creation. Furthermore, and more importantly, our model also displays effects that parallel "extensive-margin" trade diversion and trade creation when we consider the structure of matches in the economy, to which we turn in the next section.

Before doing so, we solve for closed-form expressions for equilibrium gross profits conditional on ω under preferential treatment:

$$U_t^S(\omega, t) = \frac{\alpha\left(p_w + t - \omega\right)^2}{2c - \alpha b^2}, \tag{19}$$

$$U_t^B(\omega, t) = \left[V(Q^*) - (p_w + t)Q^*\right] + \frac{2c(1-\alpha)\left(p_w + t - \omega\right)^2}{\left(2c - \alpha b^2\right)^2}. \tag{20}$$

Again, both are clearly decreasing in ω.

Hence, the gross profits of the supplier and of the buyer, when the supplier is either in *Foreign* under a PTA or in *Home*, are given by (19)-(20). The gross profits of the supplier and buyer, when the supplier is in *Foreign* without a PTA or in *ROW*, are given by (11)-(12). Note that those payoffs are the same as in (19)-(20) with $t = 0$, i.e., $U_0^S(\omega) = U_t^S(\omega, 0)$ and $U_0^B(\omega) = U_t^B(\omega, 0)$. To shorten notation, when there is little risk of confusion we drop the t argument from a function when there is no preferential treatment.

4 Structure of Matches

Initially, suppliers and buyers are not specialized toward each other. Each buyer B then matches with a supplier S in one of the three countries to form a vertical chain. We consider a Walrasian matching environment where each supplier who matches with a buyer pays a (possibly negative) fee to her buyer, and where the market for matches clears.

It is straightforward to show that matching follows a simple continuous assignment. Thus, we leave technical details to the Appendix. Importantly, Walrasian equilibrium allocations and stable outcomes coincide (Gretsky, Ostroy and Zame, 1992). That is, conditional on the equilibrium fees, no buyer or supplier could earn strictly higher profits by breaking its current match and forming a new match with a new mutually-agreeable fee. Hence, we can use the intuitive logic of stability to help describe the equilibrium.

Feasibility requires that the measure of suppliers matched not exceed the measure of available buyers (who are relatively scarce). Because all joint payoffs are strictly decreasing in ω, private efficiency requires that only the lowest-ω suppliers in each market get matched in equilibrium. Denote by $\widehat{\omega}_H$, $\widehat{\omega}_F$ and $\widehat{\omega}_{ROW}$, respectively, the hypothetical values for the cutoff levels of productivity in *Home*, *Foreign* and *ROW.* In any equilibrium, the following market-clearing condition must hold:

$$\gamma_H \int_0^{\widehat{\omega}_H} dG(\omega) + \gamma_F \int_0^{\widehat{\omega}_F} dG(\omega) + (1 - \gamma_H - \gamma_F) \int_0^{\widehat{\omega}_{ROW}} dG(\omega) = \beta. \tag{21}$$

Also, in equilibrium the marginal matches in each market must yield the same joint payoff to the members of the vertical chain. That requirement yields conditions that vary with the trade regime. We will use superscript N for cutoff values without a PTA and superscript P to denote cutoff values with a PTA.

4.1 No PTA

When there is no PTA, suppliers in *Home* enjoy the benefits from protection, while suppliers elsewhere do not. In *Foreign* and *ROW*, the joint payoff of a *B-S* chain for a given ω is equal. Hence, in equilibrium the marginal matches in *Foreign* and *ROW* must involve suppliers with the same productivity:

$$\widehat{\omega}_F^N = \widehat{\omega}_{ROW}^N. \tag{22}$$

On the other hand, a supplier with productivity ω yields a higher joint payoff if she is in *Home*, since she enjoys protection (and thus, also invests and produces more). Simple inspection of (11), (12), (19) and (20) makes clear that a vertical chain with $S(\omega)$ in *Home* yields the same joint payoff as a vertical chain with $S(\omega - t)$ in *Foreign* or *ROW*.[17] Hence,

$$\widehat{\omega}_F^N + t = \widehat{\omega}_H^N. \tag{23}$$

Collecting those conditions, it is easy to see from (21) that equilibrium without a PTA requires

$$\gamma_H G(\widetilde{\omega}_H^N) + (1 - \gamma_H) G(\widetilde{\omega}_H^N - t) = \beta. \tag{24}$$

This determines $\widetilde{\omega}_H^N(t)$—and then, through (22) and (23), also $\widetilde{\omega}_F^N(t)$ and $\widetilde{\omega}_{ROW}^N(t)$. While these cutoffs generally depend upon t, to conserve on notation we omit the t argument whenever there is no ambiguity.

Because all buyers are identical, each supplier is indifferent about the buyer to whom she is matched and cares only about the size of the fee paid. In turn, buyers care about both the size of the fee and the supplier's productivity, which affects the buyer's ultimate net profit. Equilibrium

[17] Net of the buyer's default payment $[V(Q^*) - (p_w + t)Q^*]$, the buyer and seller utilities $U_t^B(\omega, 0)$ and $U_t^S(\omega, t)$ (and their sum) are multiplicatively separable in $\{\omega, t\}$ and $\{\alpha, b, c\}$. Hence, cutoffs do not depend upon α, b and c.

is achieved when each buyer earns the same profit, so the fee must differ across matches. To see why, suppose that there was just one fee. Then a buyer matched with a relatively low-productivity supplier would earn a relatively low profit. He would prefer to match for a slightly lower fee with a higher-productivity supplier, and the higher-productivity supplier would also prefer that.

Hence, the fee paid to a buyer must depend upon the productivity of his matched supplier, as well as on whether the supplier receives a tariff preference. Specifically, without a PTA the equilibrium matching fee (M) schedule is the same for matches with suppliers in *Foreign* and *ROW* and satisfies

$$M^N_{ROW}(\omega) = M^N_F(\omega) = U^S_0(\widetilde{\omega}^N_{ROW}) - \left[U^B_0(\omega) - U^B_0(\widetilde{\omega}^N_{ROW})\right].$$

In turn, the fee schedule for matches with suppliers in *Home* satisfies

$$M^N_H(\omega, t) = U^S_t(\widetilde{\omega}^N_H, t) - \left[U^B_t(\omega, t) - U^B_t(\widetilde{\omega}^N_H, t)\right].$$

Note that all buyers earn $U^B_t(\widetilde{\omega}^N_H, t) + U^S_t(\widetilde{\omega}^N_H, t) = U^B_0(\widetilde{\omega}^N_{ROW}) + U^S_0(\widetilde{\omega}^N_{ROW}) > 0$, so their payoffs are invariant to ω. This happens because, as a higher supplier productivity raises $U^B_0(\omega)$ in *Foreign* or *ROW*, or $U^B_t(\omega, t)$ in *Home*, the buyer's fee decreases by exactly the same amount.[18] In contrast, the cutoff supplier in each market earns a payoff of exactly 0, whereas higher-productivity suppliers earn more, as they absorb the whole extra joint surplus brought about by the higher productivity through a lower fee to the buyers.

4.2 PTA

When there is a PTA between *Home* and *Foreign*, the suppliers in *Foreign* also become shielded by the (external) tariff. To characterize the equilibrium, we observe first that the market-clearing condition (21) does not change with the PTA. And once again we need to identify a condition establishing that the marginal matches in the three countries yield the same joint payoff to the members of the vertical chain. Now, however, a supplier with productivity ω generates a higher aggregate payoff if she is located in *Foreign*.

Specifically, conditions (22)-(23) now become

$$\widehat{\omega}^P_F = \widehat{\omega}^P_{ROW} + t, \tag{25}$$

and

$$\widehat{\omega}^P_F = \widehat{\omega}^P_H. \tag{26}$$

Using conditions (21), (25) and (26), we then have that equilibrium under a PTA requires

$$(\gamma_H + \gamma_F)\, G(\widetilde{\omega}^P_H) + (1 - \gamma_H - \gamma_F)\, G(\widetilde{\omega}^P_H - t) = \beta. \tag{27}$$

[18] Notice that fees from some high-productivity suppliers will be negative if $U^N_S(\widetilde{\omega}^N_{ROW}) + U^N_B(\widetilde{\omega}^N_{ROW}) < U^N_B(0)$. This is more likely to happen the higher is α.

This determines $\widetilde{\omega}_H^P$—and then, through (25) and (26), also $\widetilde{\omega}_F^P$ and $\widetilde{\omega}_{ROW}^P$.

With a PTA, the equilibrium matching fee schedule is the same for matches with suppliers in *Home* and *Foreign,* satisfying

$$M_H^P(\omega, t) = M_F^P(\omega, t) = U_t^S(\widetilde{\omega}_H^P, t) - \left[U_t^B(\omega, t) - U_t^B(\widetilde{\omega}_H^P, t)\right].$$

In turn, the fee schedule for matches with suppliers in ROW satisfies

$$M_{ROW}^P(\omega) = U_0^S(\widetilde{\omega}_{ROW}^P) - \left[U_0^B(\omega) - U_0^B(\widetilde{\omega}_{ROW}^P)\right].$$

Now, all buyers earn $U_t^B(\widetilde{\omega}_H^P, t) + U_t^S(\widetilde{\omega}_H^P, t) = U_0^B(\widetilde{\omega}_{ROW}^P) + U_0^S(\widetilde{\omega}_{ROW}^P)$. Note that this exceeds the payoff of $U_0^B(\widetilde{\omega}_{ROW}^N) + U_0^S(\widetilde{\omega}_{ROW}^N)$ that they earn under no PTA, as $\widetilde{\omega}_{ROW}^P < \widetilde{\omega}_{ROW}^N$. Once again, cutoff suppliers earn zero, while higher-productivity suppliers earn positive profits.

4.3 Comparing Equilibria under the Two Trade Regimes

We start with a useful benchmark. Define $\widetilde{\omega}_\beta$ to be such that $G(\widetilde{\omega}_\beta) = \beta$, and call this the *free-trade cutoff supplier.* When there is no tariff protection, $\widetilde{\omega}_\beta = G^{-1}(\beta)$ is the equilibrium cutoff supplier everywhere.[19] When there are more buyers, $\widetilde{\omega}_\beta$ is higher—i.e., the free-trade cutoff supplier has lower productivity. When the distribution of supplier productivity is better (in the sense of stochastic dominance), $\widetilde{\omega}_\beta$ is lower—i.e., the free-trade cutoff supplier has higher productivity.

Applying the implicit function theorem to the various equilibrium conditions, we have:

$$\widetilde{\omega}_{ROW}^P \leq \widetilde{\omega}_{ROW}^N = \widetilde{\omega}_F^N \leq \widetilde{\omega}_\beta \leq \widetilde{\omega}_F^P = \widetilde{\omega}_H^P \leq \widetilde{\omega}_H^N. \tag{28}$$

The free-trade cutoff supplier is an upper bound for cutoff suppliers in unprotected countries and is a lower bound for cutoff suppliers in protected countries. The inequalities above are strict whenever there are active suppliers in *Home*, *Foreign* and ROW.

Figure 3 shows the impact of the PTA on the equilibrium cutoffs. The dashed, upward-sloping curve is simply the 45-degree line. The other upward-sloping curve represents the equilibrium relationship between the marginal suppliers in *Home* and in ROW, which is always $\widehat{\omega}_{ROW} = \widehat{\omega}_H - t$. The relationships of these cutoffs with the marginal supplier in *Foreign* (not shown in the figure) changes with the PTA. In turn, the downward-sloped curves represent the market-clearing condition (21) under each trade regime. Without the PTA, we substitute (22) into (21); with the PTA, we substitute (26) into (21). As $\widehat{\omega}_{ROW}$ increases, more vertical chains are formed with suppliers in ROW. As a result, the number of vertical chains formed with suppliers in *Home* falls, although the rate at which it falls depends on the trade regime; under the PTA, it falls by less for a given increase in $\widehat{\omega}_{ROW}^P$, because the cutoff also falls in *Foreign.*

[19] That cutoff also obtains when all suppliers are in the same country, as well as in *Foreign* and ROW when there are no *Home* suppliers and there is no PTA, or in *Home* and *Foreign* when there are no ROW suppliers and there is a PTA.

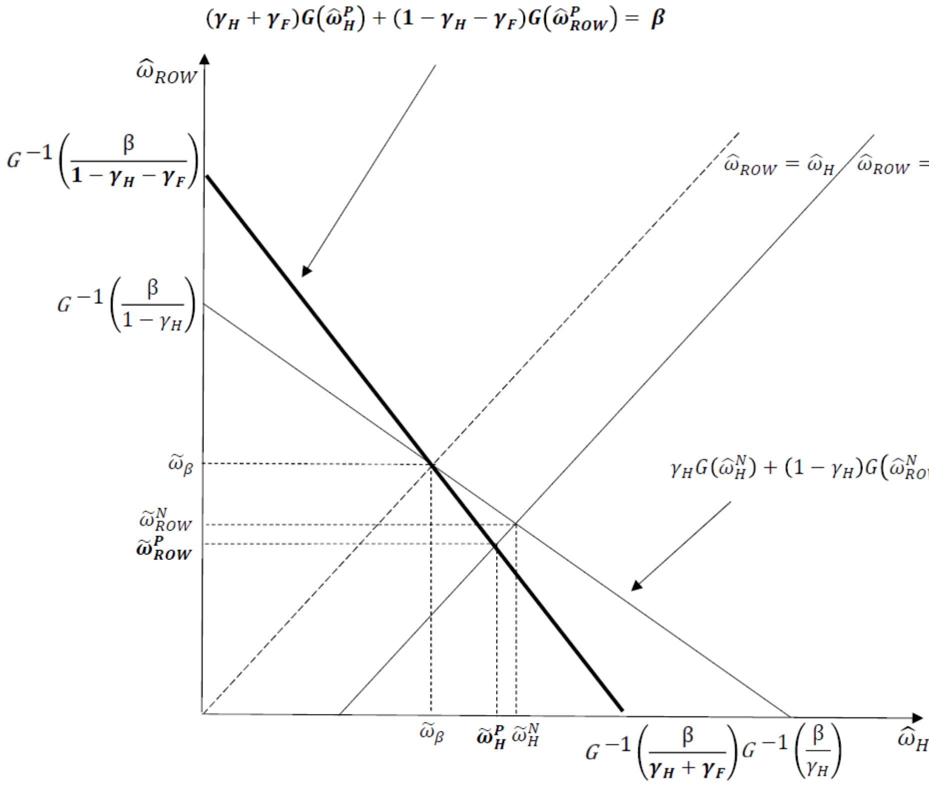

Fig. 3: Matching Equilibrium with and without the PTA

Notes: The figure shows how the PTA affects the matching equilibrium. The relationship between the cutoff *Home* and *ROW* suppliers satisfies the lower upward-sloping line, and is the same with and without a PTA. Under no PTA, the equilibrium satisfies the flatter downward-sloping curve, and the cutoff *Foreign* and *ROW* suppliers are the same, $\widetilde{\omega}_{ROW}^N$. The PTA adds discriminatory protection to mass γ_F of suppliers in *Foreign*. Equilibrium under the PTA (**bold** type) satisfies the steeper downward-sloping line, and the cutoff *Home* and *Foreign* suppliers are the same, $\widetilde{\omega}_H^P$. The figure also shows the free-trade cutoff supplier $\widetilde{\omega}_\beta$, which obtains in all countries when the tariff is zero.

As the figure illustrates, when fraction γ_F of available suppliers moves inside the trading bloc because of the PTA, the downward-sloping curve pivots around the point $(\widetilde{\omega}_\beta, \widetilde{\omega}_\beta)$ and becomes steeper, with the x-axis intercept decreasing from $G^{-1}\left(\frac{\beta}{\gamma_H}\right)$ to $G^{-1}\left(\frac{\beta}{\gamma_F+\gamma_H}\right)$ and the y-axis intercept increasing from $G^{-1}\left(\frac{\beta}{1-\gamma_H}\right)$ to $G^{-1}\left(\frac{\beta}{1-\gamma_F-\gamma_H}\right)$. As a result, the productivity of the cutoff suppliers in both *Home* and *ROW* rises with the establishment of the PTA; average supplier productivity similarly rises in those countries. By contrast, the productivity of the cutoff supplier in *Foreign,* which switches from being the same as in *ROW* to being the same as in *Home*, falls.[20] Hence, we can think of the PTA as relocating suppliers into *Foreign* from both *ROW* and *Home.*

[20] Average inverse productivity ω for *Foreign* suppliers rises as the level of the cutoff supplier rises. But note that the PTA alters investment incentives for *Foreign* suppliers, as we explore in the next section. As a result, average *observed* productivity (that is, after accounting for endogenous investment) of *Foreign* suppliers may increase or decrease with the PTA.

Comparative statics follow directly from the figure. A larger *Home* (higher β) shifts each point of the downward-sloping functions upwards, yielding higher cutoffs (and therefore lower average supplier productivity) everywhere under any trade regime. Intuitively, with more buyers, the productivity of the marginal supplier falls in all jurisdictions with and without a PTA, as buyers have to match further down in the productivity distribution. A worse supplier productivity distribution, in the sense of first-order stochastic dominance, also shifts out both downward-sloping curves and increases the level of the equilibrium cutoffs; hence, cutoff suppliers are less productive everywhere, and average supplier productivity is likewise lower.

A higher tariff t shifts curve $\widehat{\omega}_{ROW} = \widehat{\omega}_H - t$ downwards. This always increases the cutoff in *Home* and decreases the cutoff in *ROW*. However, the increase in $\widetilde{\omega}_H$ and the decrease in $\widetilde{\omega}_{ROW}$ are lower under the PTA than without it. Intuitively, a higher tariff drives a bigger wedge between the productivities of the suppliers in the marginal vertical chains. The productivity of the last supplier in *ROW* rises, while the productivity of the last supplier in *Home* falls. This effect on the marginal suppliers is smaller under the PTA because in that case $\widetilde{\omega}_F^P = \widetilde{\omega}_H^P$, so a larger mass of suppliers is affected when the cutoff in *Home* rises, while the opposite is true when $\widetilde{\omega}_{ROW}^P$ falls.

Now consider the effect of *Foreign* becoming larger. Specifically, consider that γ_F rises while γ_H falls so that $1 - \gamma_H - \gamma_F$ remains unchanged. In Figure 3, the market-clearing condition under the PTA remains the same, because it depends on the total size of the trading bloc, which does not change. As a result, the cutoffs under the PTA do not change either. On the other hand, the cutoffs without the PTA do change. This happens because the lower γ_H makes the market-clearing condition pivot around the 45-degree line and become flatter, with the y-axis intercept $G^{-1}\left(\frac{\beta}{1-\gamma_H}\right)$ falling and the x-axis intercept $G^{-1}\left(\frac{\beta}{\gamma_H}\right)$ rising. The non-PTA cutoffs $\widetilde{\omega}_H^N$ and $\widetilde{\omega}_{ROW}^N$ both rise. As a result, the magnitude of the decrease in the cutoffs in *Home* and in *ROW* induced by the PTA is greater. Intuitively, when the PTA partner is larger, the formation of the bloc induces a larger relocation of suppliers from both *Home* and *ROW*.

Importantly, none of the curves in Figure 3 are affected by the parameters of the cost function, b and c, or by supplier bargaining power, α; those parameters do not affect the matching equilibrium.

5 The Welfare Consequences of a PTA

We want to shed light on the desirability of PTAs for the world as a whole. Hence, we study how they affect *global* rather than national welfare.

We start by noting that we can express the welfare generated by a single Y-chain protected by preferential tariffs and unprotected by preferential tariffs as, respectively,[21]

$$\Psi_0(\omega) = [V(Q^*) - p_w Q^*] + p_w q_0^* - C(q_0^*, i_0^*) - I(i_0^*) \text{ and} \tag{29}$$

$$\Psi_t(\omega, t) = [V(Q^*) - p_w Q^*] + p_w q_t^* - C(q_t^*, i_t^*) - I(i_t^*). \tag{30}$$

[21] In the Appendix we develop these expressions under our cost specifications.

Observe that (29) corresponds to (7) evaluated at $(q_0^*,\, i_0^*)$. In turn, equation (30) is the equivalent expression with a tariff preference. Vertical chains located in *ROW* always yield welfare as given by (29). Vertical chains located in *Home* always yield welfare as given by (30). In contrast, the welfare generated by vertical chains located in *Foreign* depends on the trade regime.

To see the welfare changes caused by the PTA on the incumbent vertical chains in *Foreign*, observe first that the bracketed terms in expressions (29) and (30) are identical and reflect the fact that, by design, Q^* (and therefore also consumer welfare from the final good) is unchanged by the PTA. The other terms denote the surplus—including government's tariff revenue—created when a vertical chain forms, relative to the surplus B would generate if he only purchased generic inputs from *ROW*. We denote the welfare impact of giving discriminatory protection to a single vertical chain where the *Foreign* supplier has parameter ω by $\Delta\Psi(\omega,t) \equiv \Psi_t(\omega,t) - \Psi_0(\omega)$.

Aggregating over all Y-chains formed in the three countries, we have that welfare without trade agreements is given by

$$W_N = \gamma_H \int_0^{\widetilde{\omega}_H^N} \Psi_t(\omega,t)dG(\omega) + \gamma_F \int_0^{\widetilde{\omega}_F^N} \Psi_0(\omega)dG(\omega) + (1-\gamma_H-\gamma_F)\int_0^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega)dG(\omega),$$

while welfare under a PTA satisfies

$$W_P = \gamma_H \int_0^{\widetilde{\omega}_H^P} \Psi_t(\omega,t)dG(\omega) + \gamma_F \int_0^{\widetilde{\omega}_F^P} \Psi_t(\omega,t)dG(\omega) + (1-\gamma_H-\gamma_F)\int_0^{\widetilde{\omega}_{ROW}^P} \Psi_0(\omega)dG(\omega).$$

We can then express the total welfare impact of a PTA, $\Delta W(\gamma_H,\gamma_F) \equiv W_P - W_N$, as

$$\begin{aligned}\Delta W(\gamma_H,\gamma_F) = {} & \underbrace{\gamma_F \int_0^{\widetilde{\omega}_F^N} \Delta\Psi(\omega,t)dG(\omega)}_{\textit{intensive-margin welfare effect: } IM(\gamma_H,\gamma_F)} \\ & + \underbrace{\left[\gamma_F \int_{\widetilde{\omega}_F^N}^{\widetilde{\omega}_F^P} \Psi_t(\omega,t)dG(\omega) - \gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega,t)dG(\omega) - (1-\gamma_H-\gamma_F)\int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega)dG(\omega)\right]}_{\textit{extensive-margin welfare effect: } XM(\gamma_H,\gamma_F)}. \end{aligned} \quad (31)$$

The first term of (31) is the *intensive-margin welfare effect*—the welfare impact of the PTA stemming from all Y-chains with "incumbent" specialized suppliers in *Foreign* that form both with and without the PTA. We denote it by $IM(\gamma_H,\gamma_F)$.[22] The term in brackets is the *extensive-margin welfare effect*—the welfare impact due to the relocation of buyers from vertical chains with suppliers in *ROW* and in *Home* to vertical chains with suppliers in *Foreign*. We denote it by $XM(\gamma_H,\gamma_F)$.

Figure 4 illustrates those margins. To make visualization simpler, assume that buyers assortatively re-match—i.e., the buyer originally matched with the most inefficient "old" supplier (without the PTA) is matched with the least efficient "new" supplier (with the PTA). While that assumption

[22] Note that vertical chains also form with and without the PTA in *Home* $\left(\omega \in \left[0,\widetilde{\omega}_H^P\right]\right)$ and *ROW* $\left(\omega \in \left[0,\widetilde{\omega}_{ROW}^P\right]\right)$. However, the trading bloc does not change the level of tariff protection for those suppliers. Hence, nothing changes for them, and the welfare effect is nil.

is arbitrary, it is expositionally very convenient. Furthermore, as we show below, it can be used without loss of generality for the welfare analysis.

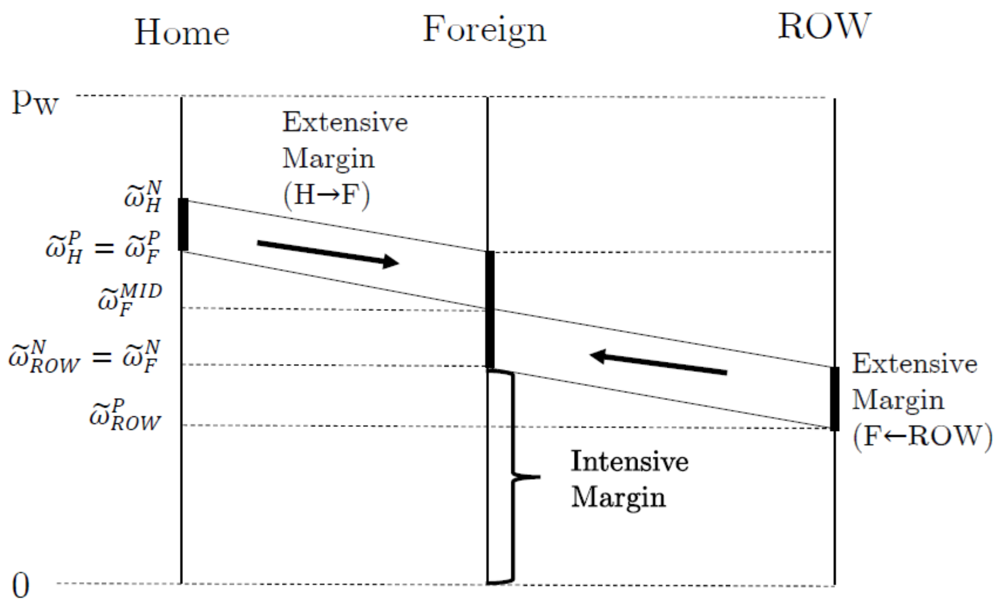

Fig. 4: Welfare Effects due to the PTA

Notes: The figure shows how the PTA relocates Y-chains across countries. Buyers break Y-chains with relatively inefficient (high-ω) suppliers in *Home* to form Y-chains with relatively efficient (low-ω) suppliers in *Foreign*. Buyers break Y-chains with relatively efficient (low-ω) suppliers in *ROW* to form Y-chains with relatively inefficient (high-ω) suppliers in *Foreign*.

As the figure shows, some incumbent suppliers in *Foreign*, $\omega \in \left[0, \widetilde{\omega}_F^N\right]$, do not enjoy discriminatory protection without the PTA but do so under a PTA. The changes in those chains define the welfare effect of the PTA due to the intensive-margin effect. Some buyers break vertical chains with relatively inefficient *Home* suppliers $\left(\omega \in \left[\widetilde{\omega}_H^P, \widetilde{\omega}_H^N\right]\right)$ and form vertical chains with relatively efficient *Foreign* suppliers $\left(\omega \in \left[\widetilde{\omega}_F^{MID}, \widetilde{\omega}_H^P\right]\right)$, where $\widetilde{\omega}_F^{MID}$ is defined so that the mass of new suppliers in *Foreign with* $\omega \geq \widetilde{\omega}_F^{MID}$ equals the mass of suppliers no longer matched in *Home*.[23] Others break vertical chains with relatively efficient *ROW* suppliers $\left(\omega \in \left[\widetilde{\omega}_{ROW}^P, \widetilde{\omega}_{ROW}^N\right]\right)$ and form vertical chains with relatively inefficient *Foreign* suppliers $\left(\omega \in \left[\widetilde{\omega}_{ROW}^N, \widetilde{\omega}_H^{MID}\right]\right)$. The combination of those relocations constitutes the extensive-margin effect.

The intensive-margin effect is relatively simple. The welfare impact for a given incumbent supplier, $\Delta\Psi(\omega, t)$, is analytically tractable and easy to characterize. The aggregate impacts for all $\omega \in [0, \widetilde{\omega}_F^N]$ are straightforward to analyze.

[23] We explicitly derive $\widetilde{\omega}_F^{MID}$ in Proposition 1. Note that this condition automatically implies that the mass of new suppliers in *Foreign with* $\omega < \widetilde{\omega}_F^{MID}$ equals the mass of suppliers no longer matched in *ROW*.

The extensive-margin effect is more complicated, because it adds surplus from new vertical chains subject to discriminatory protection (in *Foreign*), subtracts surplus from some vertical chains originally subject to discriminatory protection (in *Home*), and subtracts surplus also from some vertical chains originally not subject to discriminatory protection (in *ROW*). Furthermore, the productivities of the vertical chains that relocate from *Home* and *ROW* to *Foreign* are different. Because of the changes in all of those dimensions, analyzing the whole extensive-margin effect at once is a daunting task. However, as the next proposition shows, we can rewrite the total change in welfare isolating changes in the distribution of supplier productivity from changes in the incentives to invest and produce within each chain. As the subsequent analysis will demonstrate, this decomposition allows a much clearer view and interpretation of the several constituting parts of the total welfare effect.

Proposition 1 *There exists a* $\widetilde{\omega}_F^{MID} \in [\widetilde{\omega}_{ROW}^N, \widetilde{\omega}_H^P]$ *such that the welfare effect of a PTA can be written as the sum of an* Aggregate Within-Chain effect*,* $AWC(\gamma_H, \gamma_F)$*, a* Matching Creation effect*,* $MC(\gamma_H, \gamma_F)$*, and a* Matching Diversion effect*,* $MD(\gamma_H, \gamma_F)$*. Specifically,* $\Delta W(\gamma_H, \gamma_F) = AWC(\gamma_H, \gamma_F) + MC(\gamma_H, \gamma_F) + MD(\gamma_H, \gamma_F)$*, where*

$$\begin{cases} AWC(\gamma_H, \gamma_F) = \gamma_F \int_0^{\widetilde{\omega}_F^{MID}} \Delta\Psi(\omega, t) dG(\omega); \\ MC(\gamma_H, \gamma_F) = \left[\gamma_F \int_{\widetilde{\omega}_F^{MID}}^{\widetilde{\omega}_H^P} \Psi_t(\omega, t) dG(\omega) - \gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega, t) dG(\omega)\right]; \\ MD(\gamma_H, \gamma_F) = \left[\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega, t) dG(\omega) - (1 - \gamma_H - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega) dG(\omega)\right]. \end{cases}$$

Proof. The equilibrium conditions (24) and (27) imply that

$$\gamma_H G(\widetilde{\omega}_H^N) + (1 - \gamma_H) G(\widetilde{\omega}_{ROW}^N) = (\gamma_H + \gamma_F) G(\widetilde{\omega}_H^P) + (1 - \gamma_H - \gamma_F) G(\widetilde{\omega}_{ROW}^P).$$

This can be rewritten as

$$\gamma_F \left[G(\widetilde{\omega}_H^P) - G(\widetilde{\omega}_{ROW}^N)\right] = \gamma_H \left[G(\widetilde{\omega}_H^N) - G(\widetilde{\omega}_H^P)\right] + (1 - \gamma_H - \gamma_F) \left[G(\widetilde{\omega}_{ROW}^N) - G(\widetilde{\omega}_{ROW}^P)\right].$$

The left-hand side denotes the mass of new suppliers in *Foreign*. The right-hand side denotes the mass of old (i.e., no longer matched) suppliers, which are split between *Home* (the first term) and *ROW* (the second term). It follows that there exists a $\widetilde{\omega}_F^{MID} \in [\widetilde{\omega}_{ROW}^N, \widetilde{\omega}_H^P]$ such that

$$\gamma_H \left[G(\widetilde{\omega}_H^N) - G(\widetilde{\omega}_H^P)\right] = \gamma_F \left[G(\widetilde{\omega}_H^P) - G(\widetilde{\omega}_F^{MID})\right] \text{ and} \tag{32}$$

$$(1 - \gamma_H - \gamma_F) \left[G(\widetilde{\omega}_{ROW}^N) - G(\widetilde{\omega}_{ROW}^P)\right] = \gamma_F \left[G(\widetilde{\omega}_F^{MID}) - G(\widetilde{\omega}_{ROW}^N)\right]. \tag{33}$$

Observe that, if $\widetilde{\omega}_F^{MID}$ satisfies equation (32), it also satisfies equation (33)—and vice versa.

Now use the fact that $\widetilde{\omega}_F^N = \widetilde{\omega}_{ROW}^N$ and $\widetilde{\omega}_F^P = \widetilde{\omega}_H^P$ to rewrite the first term of $XM(\gamma_H, \gamma_F)$ in (31) as $\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_H^P} \Psi_t(\omega, t) dG(\omega)$. Moreover, it is straightforward to divide up that integral using

$\widetilde{\omega}_F^{MID}$, and then rearrange the terms in (31) as

$$\begin{aligned}\Delta W(\gamma_H,\gamma_F) &= \gamma_F \int_0^{\widetilde{\omega}_F^N} \Delta\Psi(\omega,t)dG(\omega) \\ &+ \left[\gamma_F \int_{\widetilde{\omega}_F^{MID}}^{\widetilde{\omega}_H^P} \Psi_t(\omega,t)dG(\omega) - \gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega,t)dG(\omega)\right] \\ &+ \left[\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_t(\omega,t)dG(\omega) - (1-\gamma_H-\gamma_F) \int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega)dG(\omega)\right].\end{aligned} \tag{34}$$

The first term in brackets is what we denote as $MC(\gamma_H,\gamma_F)$.

Now add and subtract $\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega,t)dG(\omega)$ to equation (34). Since $\widetilde{\omega}_F^N = \widetilde{\omega}_{ROW}^N$, the lower limit of this integral is equal to the upper limit of the first integral in (34). We can then add $\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Delta\Psi(\omega,t)dG(\omega) = \gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_t(\omega,t)dG(\omega) - \gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega)dG(\omega)$ to the first integral in (34), which then becomes $\gamma_F \int_0^{\widetilde{\omega}_F^{MID}} \Delta\Psi(\omega,t)dG(\omega)$. This is what we denote as $AWC(\gamma_H,\gamma_F)$. With the residual terms, the second square bracket in (34) becomes $[\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega)dG(\omega) - (1 - \gamma_H - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega)dG(\omega)]$, which is what we denote as $MD(\gamma_H,\gamma_F)$.

Finally, by the definition of $\widetilde{\omega}_F^{MID}$, we know that the "new" and "old" suppliers in the MC and MD terms have each the same probability mass of suppliers. Therefore, we can write the total welfare of a PTA as $\Delta W(\gamma_H,\gamma_F) = AWC(\gamma_H,\gamma_F) + MC(\gamma_H,\gamma_F) + MD(\gamma_H,\gamma_F)$. ■

The welfare decomposition in Proposition 1 allows us to investigate separately the welfare consequences due to, on one hand, altered incentives to invest and produce for a given set of producers, and on the other hand, changes in the distribution of productivity of existing suppliers. Accordingly, we study each of them in turn. However, to get to the aggregate effects, we first analyze the welfare consequences of a PTA due to a given incumbent supplier in *Foreign* whose productivity parameter ω is arbitrary (subsection 5.1). In subsection 5.2 we study the aggregate within-chain effect and then, in subsection 5.3, we analyze the matching creation and matching diversion effects. We consider the total effect in section 6.

5.1 A Single Incumbent Y-Chain

Within a given Y-chain with an incumbent supplier in *Foreign*, a PTA induces an increase in the sourcing of specialized inputs, coupled with changes in the cost of producing them and an increase in the cost of investment incurred by S. The total welfare effect for a single Y-chain can be expressed under our cost specifications as (see the Appendix for the expressions for Ψ_0 and Ψ_t):

$$\Delta\Psi(\omega,t) = \frac{t}{(2c-\alpha b^2)^2}\left[2b^2\alpha(1-\alpha)(p_w-\omega) - t(2c+\alpha^2 b^2 - 2\alpha b^2)\right]. \tag{35}$$

As this function is linear in inverse productivity ω and quadratic in the external tariff t, it permits a straightforward comparative static analysis of these and all other parameters.

However, we gain deeper insights by first splitting $\Delta\Psi(\omega, t)$ into two conceptual effects, *relationship strengthening* ($\Delta\Psi_{RS}$) and *sourcing diversion* ($\Delta\Psi_{SD}$), with $\Delta\Psi(\omega, t) = \Delta\Psi_{RS} + \Delta\Psi_{SD}$. The *relationship-strengthening effect* reflects the welfare consequences of the PTA on the (ex-ante) investment decisions while assuming that, given the investment, the (ex-post) sourcing decision would be socially efficient. It corresponds to the additional surplus created by S's extra investment on the production of q_1^*—i.e., the reduction in specialized input cost relative to the cost of using generic inputs in the production of the ex-post socially efficient level q_1^*, illustrated by areas $\mathbf{C} + \mathbf{D}$ in Figure 2—net of the increased investment cost. Specifically,

$$\Delta\Psi_{RS} = p_w(q_1^* - q_0^*) + [C(q_0^*, i_0^*) - C(q_1^*, i_t^*)] - [I(i_t^*) - I(i_0^*)]. \tag{36}$$

After some manipulation, this expression can be rewritten as

$$\Delta\Psi_{RS} = \frac{2c - b^2}{2c}\Delta i\,(HUP_0 - EXC_t)\,. \tag{37}$$

Expression (37) is very intuitive. There is underinvestment in the absence of trade agreements ($HUP_0 > 0$), and the increase in investment ($\Delta i > 0$) mitigates that original inefficiency. The first term in the parenthesis reflects the ensuing welfare gains from moving the supplier's investment toward the first-best level. However, Δi may be too large and yield overinvestment under a PTA, in which case $EXC_t > 0$. The second term in the parenthesis reflects the welfare losses from inducing the supplier to invest above the first-best level. The sign of $\Delta\Psi_{RS}$ is positive as long as the PTA moves investment *closer* to the efficient level. Naturally, if the underinvestment problem remains under the PTA despite the extra investment, then $EXC_t < 0$ and $\Delta\Psi_{RS} > 0$ for sure.

It also follows from expression (37) that $\Delta\Psi_{RS}$ is non-monotonic in Δi. When Δi is low, the relationship-strengthening effect is positive and increasing in Δi. But when Δi is very high, $HUP_0 - EXC_t < 0$ and an increase in Δi amplifies the distortion in investment spending. If investment is completely unresponsive to investment ($b = 0$), the relationship-strengthening effect vanishes.

In turn, the *sourcing-diversion effect* reflects the welfare consequences of the PTA due to distortions in sourcing decisions—i.e. the deadweight loss from using specialized inputs that are too costly—given the investment choice under the PTA. This is the direct result of the protection the tariff preference affords S by skewing the sourcing decision away from generic inputs. Explicitly,

$$\begin{aligned}\Delta\Psi_{SD} &= C(q_1^*, i_t^*) - C(q_t^*, i_t^*) + p_w(q_t^* - q_1^*) \\ &= -\frac{t^2}{2c}.\end{aligned} \tag{38}$$

This corresponds to (the negative of) area $\mathbf{E}$ in Figure 2. It does not depend upon investment responsiveness, b.

A single Y-chain generates higher welfare under a PTA provided that the relationship-strengthening effect is positive and dominates the sourcing diversion effect, i.e., $\Delta\Psi_{RS} \geq |\Delta\Psi_{SD}|$. This comparison highlights a trade-off between improvements in the efficiency of the production process ($\Delta\Psi_{RS}$) and tariff-induced allocative inefficiency ($\Delta\Psi_{SD}$).

A key determinant of the balance of this trade-off is the supplier's (inverse) productivity parameter, ω, which shifts her marginal cost function. From Lemmas 1 and 2 we have that $\frac{\partial HUP_0}{\partial \omega} < 0$ and $\frac{\partial EXC_t}{\partial \omega} > 0$. Hence, the potential efficiency-enhancing aspect of a PTA is unambiguously more important for more productive suppliers (which have a lower ω).

The sourcing-diversion effect, on the other hand, does *not* change with ω. Since neither the level of productivity nor investment affects the slope of the marginal cost curve, the implied deadweight loss is a constant function of both. The upshot is that, for a given incumbent Y-chain in *Foreign*, the downside of a PTA is unaffected by the productivity of the specialized supplier, whereas the upside rises with it. The next result follows directly from those observations.

Lemma 3 *Higher supplier productivity induces a stronger relationship-strengthening effect, but has no impact on the sourcing-diversion effect of a PTA. Hence, the welfare effect for a single incumbent Y-chain $\Delta\Psi(\omega, t)$ is decreasing in ω.*

Next, consider the effects of the preferential margin, t. It affects $\Delta\Psi_{RS}$ in almost the exact same non-monotonic way as Δi. By increasing investment, it initially increases $\Delta\Psi_{RS}$. But the positive effect diminishes with t, and for sufficiently high t, investment becomes excessive and $\Delta\Psi_{RS}$ turns negative. The absolute value of the sourcing-diversion effect, $|\Delta\Psi_{SD}|$, increases in t, and at an increasing rate.

Hence, if t is sufficiently low, the first-order gain from the relationship-strengthening effect dominates the second-order loss from the sourcing-diversion effect within an incumbent Y-chain. In that case, the net within-chain impact of a PTA is necessarily positive. For higher t, however, the sourcing-diversion effect dominates, and the within-chain impact is negative. Indeed, $\Delta\Psi(\omega, t)$ is an inverted-U function of t, achieves a (positive) maximum for a unique t, and is negative for t sufficiently high. The next lemma proves those claims.

Lemma 4 *The relationship-strengthening effect is an inverted-U function of the external tariff, while the sourcing-diversion effect is a strictly decreasing function of the external tariff. The overall within-chain welfare impact of a PTA, $\Delta\Psi(\omega, t)$, is also an inverted-U function of t. For $\omega \in [0, p_w)$, it achieves a maximum at*

$$\widehat{t}(\omega) \equiv \frac{\alpha(1-\alpha)b^2(p_w - \omega)}{2c - 2\alpha b^2 + \alpha^2 b^2}. \tag{39}$$

Furthermore, $\Delta\Psi(\omega, t) > 0$ if $t \in \left(0, 2\widehat{t}(\omega)\right)$ and $\Delta\Psi(\omega, t) < 0$ if $t > 2\widehat{t}(\omega)$.

Note that the within-chain welfare-maximizing tariff depends on the productivity of the supplier, so we can define $\widehat{t}(\omega)$ for each ω. In particular, $\widehat{t}(\omega)$ is higher when ω is lower. Thus, when the supplier is more productive, the external tariff that maximizes the welfare effect of the PTA stemming from that Y-chain is higher, being highest for the lowest supplier, $\omega = 0$:

$$\widehat{t}(0) \equiv \frac{\alpha(1-\alpha)b^2 p_w}{2c - 2\alpha b^2 + \alpha^2 b^2}.$$

Intuitively, a lower ω makes the hold-up problem more severe, so a larger investment boost resulting from a higher preferential margin under the PTA generates a larger welfare gain. Similarly, the range of external tariffs under which $\Delta\Psi(\omega, t) > 0$ expands as ω falls. An important implication of Lemma 4 is that, for a tariff $t < 2\widehat{t}(0)$, the PTA raises welfare provided that ω is sufficiently low. In contrast, the tariff may be so high ($t > 2\widehat{t}(0)$) that the welfare effect is negative for any ω.

Lemma 4 places some additional bounds on the benefits of a PTA stemming from the relationship-strengthening effect. If suppliers' bargaining power α is either very high or very low, the potential for the PTA to raise welfare is severely limited, in the sense of reducing $2\widehat{t}(\omega)$. When α is very low, the PTA is ineffective at stimulating additional investment. When α is very high, the under-investment problem without a PTA is minimal, and the PTA yields over-investment. A very low b, which implies that investment is relatively unproductive, also yields a tight $2\widehat{t}(\omega)$ bound.

5.2 The Aggregate Within-Chain Effect

The aggregate within-chain welfare effect sums the relationship-strengthening and the sourcing-diversion effects of the PTA across all vertical chains with $\omega \in \left(0, \widetilde{\omega}_F^{MID}\right)$. Hence,

$$AWC(\gamma_H, \gamma_F) = \gamma_F \int_0^{\widetilde{\omega}_F^{MID}} \Delta\Psi(\omega, t) dG(\omega). \tag{40}$$

Given Lemmas 3 and 4, it is easy to analyze $AWC(\gamma_H, \gamma_F)$. Lemma 3 implies that if the PTA increases welfare stemming from supplier ω', then it also increases welfare due to all chains with suppliers $\omega \in [0, \omega')$. And if the PTA decreases welfare stemming from supplier ω', then it also decreases welfare due to all chains with suppliers $\omega \in (\omega', p_w]$.

Lemma 4 further implies that if the welfare effect from the most productive incumbent supplier is negative, $\Delta\Psi(0, t) \leq 0$, as obtains for any $t \geq 2\widehat{t}(0)$, then $AWC < 0$. Similarly, if the PTA increases welfare from the vertical chain with productivity $\widetilde{\omega}_F^{MID}$, $\Delta\Psi(\widetilde{\omega}_F^{MID}, t) \geq 0$, then $AWC > 0$. Because $\widetilde{\omega}_F^{MID} \leq \widetilde{\omega}_\beta + t$, the tariff that generates zero welfare effect for the maximum cutoff supplier, $\widetilde{\omega}_\beta + t$, guarantees a positive AWC. The following proposition summarizes these results.

Proposition 2 *If $t \leq t_{AWC} \equiv \frac{2b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_\beta)}{2c - \alpha^2 b^2}$, then the aggregate within-chain effect is positive. If $t \geq 2\widehat{t}(0)$, then the aggregate within-chain effect is negative.*

If the tariff is between t_{AWC} and $2\widehat{t}(0)$, then $\Delta\Psi(0, t) > 0$ but $\Delta\Psi(\widetilde{\omega}_\beta + t, t) < 0$, and the sign of $AWC(\gamma_H, \gamma_F)$ depends on γ_H, γ_F, and $G(\omega)$.

5.3 Matching Creation and Matching Diversion

The matching-creation effect is defined as

$$MC(\gamma_H, \gamma_F) = \gamma_F \int_{\widetilde{\omega}_F^{MID}}^{\widetilde{\omega}_H^P} \Psi_t(\omega, t) dG(\omega) - \gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega, t) dG(\omega).$$

It shows the change in surplus resulting from buyers abandoning matches with relatively inefficient *Home* suppliers $\left(\omega \in \left[\widetilde{\omega}_H^P, \widetilde{\omega}_H^N\right]\right)$ and rematching with relatively efficient *Foreign* suppliers $\left(\omega \in \left[\widetilde{\omega}_F^{MID}, \widetilde{\omega}_H^P\right]\right)$. The following proposition shows that this effect is unambiguously positive.

Proposition 3 *The matching-creation effect is positive:* $MC(\gamma_H, \gamma_F) > 0$.

Proof. Because $\Psi_t(\omega, t)$ is strictly decreasing in ω and G is continuous and strictly increasing in ω, it follows that

$$\begin{aligned}
\gamma_F \int_{\widetilde{\omega}_F^{MID}}^{\widetilde{\omega}_H^P} \Psi_t(\omega, t) dG(\omega) &> \left[\gamma_F \int_{\widetilde{\omega}_F^{MID}}^{\widetilde{\omega}_H^P} dG(\omega)\right] \Psi_t(\widetilde{\omega}_H^P, t) \\
&= \left[\gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} dG(\omega)\right] \Psi_t(\widetilde{\omega}_H^P, t) > \gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega, t) dG(\omega),
\end{aligned}$$

where the equality uses the definition of $\widetilde{\omega}_F^{MID}$ from equation (32). This completes the proof. ■

The intuition is simple. The construction of the MC effect isolates a mass of buyers whose size is $\gamma_H \left[G(\widetilde{\omega}_H^N(t)) - G(\widetilde{\omega}_H^P(t))\right] = \gamma_F \left[G(\widetilde{\omega}_H^P(t)) - G(\widetilde{\omega}_F^{MID})\right]$. This mass matches with *Home* suppliers without a PTA and *Foreign* suppliers under a PTA. For both groups of Y-chains, there is discriminatory protection. This means that the only difference between the groups is that the new suppliers in *Foreign* have higher productivity than the old suppliers in *Home.* Hence, this relocation surely enhances global welfare.

In turn, the matching-diversion effect is defined as

$$MD(\gamma_H, \gamma_F) = \gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega) dG(\omega) - (1 - \gamma_H - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} \Psi_0(\omega) dG(\omega).$$

It shows the change in surplus resulting from buyers abandoning matches with relatively efficient ROW suppliers $\left(\omega \in \left[\widetilde{\omega}_{ROW}^P, \widetilde{\omega}_{ROW}^N\right]\right)$ and rematching with relatively inefficient *Foreign* suppliers $\left(\omega \in \left[\widetilde{\omega}_{ROW}^N, \widetilde{\omega}_H^{MID}\right]\right)$. The following proposition shows that this effect is unambiguously negative.

Proposition 4 *The matching-diversion effect is negative:* $MD(\gamma_H, \gamma_F) < 0$.

The intuition and proof are analogous to that of Proposition 3. The construction of the MD effect isolates a mass of buyers whose size is $(1 - \gamma_H - \gamma_F) \left[G(\widetilde{\omega}_{ROW}^N(t)) - G(\widetilde{\omega}_{ROW}^P(t))\right] = \gamma_F \left[G(\widetilde{\omega}_F^{MID}) - G(\widetilde{\omega}_{ROW}^N(t))\right]$. This mass matches with ROW suppliers without a PTA and *Foreign* suppliers under a PTA. The former group does not enjoy discriminatory protection, while the latter does—and that is why there is such a relocation. However, the decomposition from Proposition 1 allows us to treat them *as if* neither group of Y-chains had discriminatory protection. That means that the only difference between the groups is that the new suppliers in *Foreign* have lower productivity than the old suppliers in ROW. Hence, when deprived of the effects of discriminatory protection, this inefficient relocation of suppliers across vertical chains necessarily lowers global welfare.

In general, one cannot say whether matching creation dominates or is dominated by matching diversion, as the comparison between the two effects depends on the parameters of the model and on $G(\omega)$. Figure 5 illustrates how the two effects vary with the share of specialized suppliers in *Home*, γ_H (for γ_F fixed at 0.5) when ω is distributed uniformly in the range $[0, p_w]$. In one extreme, when $\gamma_H = 0$, MC is shut down: there are no (relatively inefficient) suppliers to relocate from *Home*. In the other extreme, when $\gamma_H = 0.5$ (and therefore $1 - \gamma_H - \gamma_F = 0$), MD is shut down: there are no (relatively efficient) suppliers to relocate from ROW. Moreover, MC increases while $|MD|$ decreases monotonically in γ_H, as the scope for beneficial relocations from *Home* to *Foreign* enlarges and the scope for harmful relocations from ROW to *Foreign* shrinks.

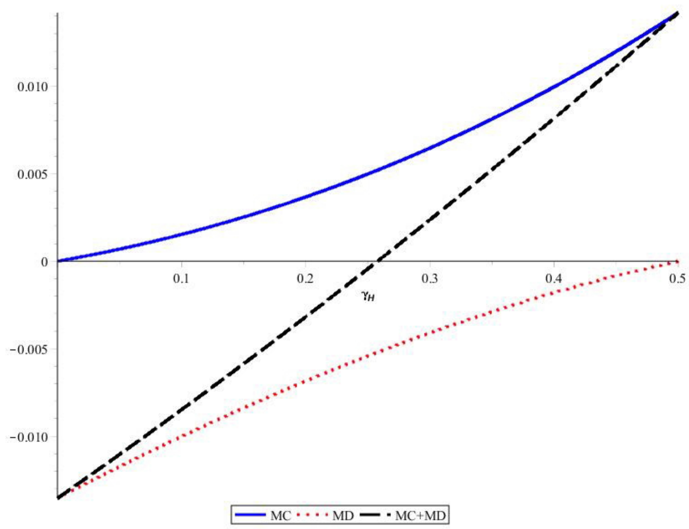

Fig. 5: Matching-Creation and Matching-Diversion Effects, Uniform Distribution

Notes: This figure shows the matching-creation and matching-diversion effects when ω is distributed uniformly on $[0, p_w]$ and half of all suppliers are located in *Foreign* ($\gamma_F = 0.5$). The values for the other parameters are $p_w = c = 1$, $b = 1.25$, $t = 0.1$ and $\alpha = \beta = 0.5$.

Now, if we leave the distribution of productivity unrestricted, then such a monotonicity does not need to hold. Nevertheless, one can generate additional results on the relationship between matching creation and matching diversion conditional on properties of the distribution function. For example, if $G(\omega)$ were convex, as in a Pareto distribution, then $|MD|$ would tend to increase while MC would tend to decrease, relative to their levels under the uniform distribution. The reason is that, with a convex distribution, the density of suppliers decreases with productivity (i.e., increases with ω). This implies that, for a given measure of suppliers relocated from ROW, a wider range of ω is affected, meaning that some suppliers with very high productivity lose their matches with the PTA. Conversely, in *Home* a narrower range of ω is affected, meaning that only suppliers with not-so-low productivity are relocated from *Home*. Such an imbalance between MC and MD tends to increase as $G(\omega)$ becomes more convex, i.e., as the distribution of productivity becomes more skewed.

6 Overall Welfare Effects

We consider now the overall welfare effect of a PTA under global sourcing. Unsurprisingly, it is generally ambiguous. However, we know that matching creation always raises welfare and that matching diversion always lowers welfare. The aggregate within-chain effect is generally ambiguous, but we know that within-chain welfare effects are positive for low levels of discriminatory protection. It turns out that, for very low levels of the external tariff, that effect is first-order while effects from supplier relocations are second-order.

Proposition 5 *The overall welfare effect of a PTA is positive for sufficiently low t.*

An immediate implication of Proposition 5 is that, if the PTA lowers aggregate welfare, it is because the external tariff—a policy variable that could also be changed with the agreement—is too high. More generally, although the welfare impact of a PTA is ambiguous, it will tend to one direction or the other under specific circumstances. To highlight that, we analyze some benchmark cases.

Consider first the situation where $\gamma_H + \gamma_F = 1$. In that case, since there are no specialized suppliers in ROW matching with *Home* buyers, there is no matching-diversion effect. We can think of that as a situation where the PTA members are strong "natural partners," perhaps due to geographical remoteness, as for example Australia and New Zealand.

In subsection 6.2, we then consider the situation where $\gamma_H = 0$. In that case, there is no matching-creation effect. We can think of that as a situation where *Home* specializes in headquarters services, as in a North-South model. Since by design we do not have intensive-margin trade creation, the North-South case stacks the odds against making the overall welfare effect positive.

A useful special case that fits in both frameworks is when $\gamma_F = 1$. This can be interpreted as the limiting situation where the preferential partner is very large and has a North-South type relationship, e.g., the US for Mexico within NAFTA. Analytically, setting $\gamma_F = 1$ allows us to keep the set of vertical chains unchanged by the PTA—so that welfare changes only because of the intensive margin. We discuss this case in detail in the next subsection and return to it at the end of subsection 6.2.

6.1 The Natural Partners Case ($\gamma_H + \gamma_F = 1$)

When there are no ROW specialized suppliers, equilibrium matching is straightforward. Under no PTA, the cutoff suppliers satisfy $\widetilde{\omega}_H^N = \widetilde{\omega}_F^N + t$ and $\gamma_H G(\widetilde{\omega}_H^N) + \gamma_F G(\widetilde{\omega}_F^N) = \beta$. With the PTA, the free-trade cutoff supplier $\widetilde{\omega}_\beta$ obtains in both *Home* and *Foreign*. For any t, the cutoff suppliers satisfy $\widetilde{\omega}_F^N \leq \widetilde{\omega}_\beta \leq \widetilde{\omega}_H^N$. Because there are no ROW suppliers, all supplier relocations flow from *Home* to *Foreign* and there is no matching-diversion effect. Furthermore $\widetilde{\omega}_F^{MID} = \widetilde{\omega}_F^N$, as follows from (32), and the aggregate within-chain effect equals the intensive-margin welfare effect.

We can then write

$$\begin{aligned}
&\Delta W(\gamma_H, 1-\gamma_H) = AWC(\gamma_H, 1-\gamma_H) + MC(\gamma_H, 1-\gamma_H) \\
&= (1-\gamma_H)\int_0^{\widetilde{\omega}_F^N} \Delta\Psi(\omega,t)dG(\omega) + \left[(1-\gamma_H)\int_{\widetilde{\omega}_F^N}^{\widetilde{\omega}_\beta} \Psi_t(\omega,t)dG(\omega) - \gamma_H \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_H^N} \Psi_t(\omega,t)dG(\omega)\right].
\end{aligned}$$

Generally, the direction of this welfare effect is ambiguous, as the sign of the intensive-margin effect depends upon the cutoff supplier under no PTA, $\widetilde{\omega}_F^N$. But because the matching-creation effect is positive, the total welfare effect exceeds the intensive-margin effect. Because all intensive-margin suppliers are at least as productive as the free-trade cutoff $\widetilde{\omega}_\beta$, Lemma 4 implies the following sufficient condition on tariffs for a positive overall welfare effect.

Proposition 6 *For natural partners ($\gamma_H + \gamma_F = 1$), if $t \leq 2\widehat{t}(\widetilde{\omega}_\beta)$, then the welfare effect of the PTA is strictly positive.*

Hence, the welfare effect of a PTA in the natural partners case hinges on the intensive-margin effect. To study it more carefully, let us consider the special case of a "large" natural trading partner ($\gamma_F = 1$), where the welfare effect of the PTA is just the intensive-margin effect evaluated at $\gamma_F = 1$. As the free-trade cutoff supplier $\widetilde{\omega}_\beta$ obtains with and without a PTA, we have

$$\Delta W(0,1) = AWC(0,1) = IM(0,1) = \int_0^{\widetilde{\omega}_\beta} \Delta\Psi(\omega,t)dG(\omega).$$

The total welfare effect is then an inverted-U function of t driven by the (aggregated) tradeoffs between relationship-strengthening and sourcing-diversion effects.[24]

Proposition 7 *In a PTA with a large ($\gamma_F = 1$), natural trading partner, the aggregate within-chain effect corresponds to the intensive-margin effect and to the total welfare impact of a PTA. It is an inverted-U function of t, achieving a maximum at*

$$\widehat{t}_W^{\gamma_F=1} \equiv \frac{\alpha(1-\alpha)b^2\left[p_w - \mathbf{E}\left(\omega; \omega \leq \widetilde{\omega}_\beta\right)\right]}{2c - 2\alpha b^2 + \alpha^2 b^2} \in \left(\widehat{t}(\widetilde{\omega}_\beta), \widehat{t}(0)\right), \tag{41}$$

and being positive if and only if $t \leq 2\widehat{t}_W^{\gamma_F=1}$.

For a large, natural trading partner, there is a level of preferential margin $\widehat{t}_W^{\gamma_F=1}$ that optimally trades off the gains from relationship strengthening against the losses from sourcing diversion.[25]

[24] Observe that, if $\gamma_H \neq 0$, the aggregate within-chain effect may not be an inverted-U with respect to t. The reason is that, in that case, the cutoff supplier depends upon the tariff. As a result, first- and second-order effects of t depend upon the densities at the cutoffs in both *Home* and *Foreign*.

[25] Observe that $\mathbf{E}\left(\omega; \omega \leq \widetilde{\omega}_\beta\right)$ is fully determined by the distribution of ω and by parameter β, so $\widehat{t}_W^{\gamma_F=1}$ is a function of primitives only.

This optimum is higher than the optimum for the free-trade cutoff supplier, $\widehat{t}(\widetilde{\omega}_\beta)$, and is higher when the average productivity of active suppliers is higher (i.e., when $\mathbf{E}(\omega; \omega \leq \widetilde{\omega}_\beta)$ is lower). The reason is that, when suppliers are more productive, the original hold-up problem is more severe (Lemma 1), so it pays (from a social perspective) to have a larger margin of preference to boost the relationship-strengthening effect. The same factors that determine $\widehat{t}_W^{\gamma_F=1}$ also determine the highest level of preferential margin under which the welfare effect is positive, $2\widehat{t}_W^{\gamma_F=1}$.

Notice that none of the cutoff suppliers depend on either suppliers' bargaining power (α) or the productivity of investment (b). Therefore, the comparative statics with respect to them for a single vertical chain, discussed at the end of subsection 5.1, apply here as well.

Furthermore, using Lemma 3 we can precisely characterize the effect of the distribution of productivity. In particular, let us say that $G_2(\omega)$ **FOSD** $G_1(\omega)$ when distribution $G_2(\omega)$ first-order stochastically dominates distribution $G_1(\omega)$. In that case, the welfare effect of the PTA is unambiguously higher under $G_1(\omega)$ than under $G_2(\omega)$.

Proposition 8 *In a PTA with a large, natural trading partner ($\gamma_F = 1$), if $G_2(\omega)$ **FOSD** $G_1(\omega)$, then $\Delta W(0, 1; G_1) > \Delta W(0, 1; G_2)$.*

Proposition 8 implies that, in the context of a PTA with a large natural partner, welfare rises provided that the distribution of suppliers has sufficiently high productivity, but not otherwise. A corollary is that, if one could identify a distribution $G_0(\omega)$ such that $\Delta W(0, 1; G_0) = 0$, one would know that $\Delta W(0, 1; G) > 0$ under all distribution functions that are "better" than $G_0(\omega)$, in the sense of being first-order stochastically dominated by $G_0(\omega)$, and $\Delta W(0, 1; G) < 0$ under all distribution functions with the opposite property. As such, Proposition 8 may be used as a guide for industry exclusion within a PTA. If one could rank industries within a PTA using a **FOSD** criterion (which should generally be related to measures of comparative advantage), then an "optimal exclusion" criterion would indicate that all industries with an inverse productivity distribution that **FOSD** $G_0(\omega)$ should be excluded from the agreement, whereas all industries with an inverse productivity distribution first-order stochastically dominated by $G_0(\omega)$ should be integral parts of it.

The following example illustrates Propositions 7 and 8.

Example 1 *Let $\gamma_F = 1$ and consider that productivity $1/\omega$ follows a Pareto distribution with lower distribution bound $1/p_w$ and shape parameter $k \geq 1$. This yields $G(\omega) = \left(\frac{\omega}{p_w}\right)^k$ for $\omega \in [0, p_w]$. Consider then the distributions for $k = 1, 2$, $G_{k1}(\omega) = \frac{\omega}{p_w}$ and $G_{k2}(\omega) = \left(\frac{\omega}{p_w}\right)^2$. $G_{k1}(\omega)$ corresponds to a uniform distribution. Clearly, $G_{k2}(\omega)$ **FOSD** $G_{k1}(\omega)$. Equilibrium cutoffs are $\widetilde{\omega}_{\beta 1} = \beta p_w$ and $\widetilde{\omega}_{\beta 2} = \sqrt{\beta} p_w$, and $\mathbf{E}(\omega; \omega \leq \widetilde{\omega}_{\beta 1}) < \mathbf{E}(\omega; \omega \leq \widetilde{\omega}_{\beta 2})$. Figure 6 shows ΔW for each of them as a function of the tariff. Following Proposition 8, ΔW is higher for every t under $G_{k1}(\omega)$. Following Proposition 7, for both distributions ΔW is an inverted-U with respect to t, is strictly positive for low external tariffs, and is strictly negative for tariffs more than twice as high as the tariff that maximizes it. Furthermore, the peak of ΔW obtains for a higher t under $G_{k1}(\omega)$.*

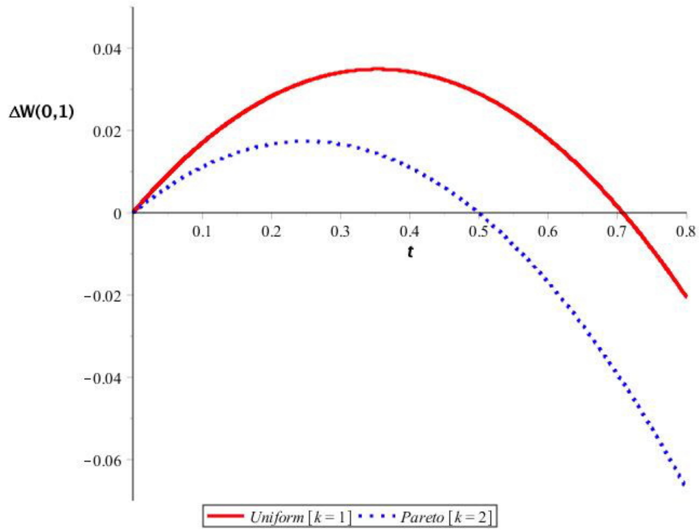

Fig. 6: Welfare Effects of the PTA under Pareto and Uniform Distributions

Notes: This figure shows the welfare effect of the PTA when all suppliers are located in *Foreign* and inverse productivity $\frac{1}{\omega}$ follows a Pareto distribution with shape parameter $k \in \{1, 2\}$. The other parameters are the same as in Figure 5.

6.2 The North-South Case ($\gamma_H = 0$)

Now consider the case without *Home* suppliers. As this specification shuts down the matching-creation effect, it enables a tractable and detailed analysis of the welfare effects coming from the $ROW \to$*Foreign* extensive margin.

It is straightforward to characterize matching. Without the PTA, the free-trade cutoff supplier $\widetilde{\omega}_\beta$ obtains in both *Foreign* and ROW. With the PTA, the cutoff suppliers satisfy $\widetilde{\omega}_F^P = \widetilde{\omega}_{ROW}^P + t$ and $\gamma_F G(\widetilde{\omega}_F^P) + (1 - \gamma_F) G(\widetilde{\omega}_{ROW}^P) = \beta$. It follows from equation (32) that $\widetilde{\omega}_F^{MID} = \widetilde{\omega}_F^P$.

Recalling that $\Delta W(0, \gamma_F) = IM(\gamma_F) + XM(0, \gamma_F)$, we can write

$$XM(0, \gamma_F) = \gamma_F \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_F^P} \Delta\Psi(\omega, t) dG(\omega) + \left[\gamma_F \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_F^P} \Psi_0(\omega) dG(\omega) - (1 - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_\beta} \Psi_0(\omega) dG(\omega) \right]. \tag{42}$$

The term in brackets corresponds to matching diversion when $\gamma_H = 0$. Accordingly, the difference between $XM(0, \gamma_F)$ and $MD(0, \gamma_F)$ is the first term in (42), due to the change in investment and production of the new suppliers in *Foreign*. These changes may generate welfare benefits, and those benefits may outweigh the matching-diversion effect, as we illustrate later in this subsection. However, we will see that this can occur only under fairly special conditions—tariffs need to be low and the density of suppliers needs to be such that the magnitude of the matching-diversion effect is also low.

To analyze XM, we delve deeper into the mechanics of supplier relocation. There is a relocation of buyers from vertical chains with ROW suppliers ($\omega \in [\widetilde{\omega}^P_{ROW}(t), \widetilde{\omega}_\beta]$) to vertical chains with *Foreign* suppliers ($\omega \in [\widetilde{\omega}_\beta, \widetilde{\omega}^P_F(t)]$). For a small change in the tariff from t to $t + dt$, the cutoff supplier $\widetilde{\omega}^P_{ROW}(t)$ falls, the cutoff supplier $\widetilde{\omega}^P_F(t)$ rises, and an additional number of supplier relocations occur. The exact measure of relocations induced by the increase dt is a function of both the density of cutoff suppliers in *Foreign*, $\gamma_F g(\widetilde{\omega}^P_F(t))$, and the density of cutoff suppliers in ROW, $(1 - \gamma_F) g(\widetilde{\omega}_{ROW}(t))$.

We call this measure the *flow rate* of relocations. We can derive a precise expression for this flow rate using a change of variables to rewrite the extensive-margin welfare effect as[26]

$$XM(0, \gamma_F) = \int_0^t \left[\Psi_t(\widetilde{\omega}^P_F(x), t) - \Psi_0(\widetilde{\omega}^P_{ROW}(x))\right] \phi(x; \gamma_F, G) dx. \tag{43}$$

The new argument x is a hypothetical tariff that affects only the (monotonic) cutoffs $\widetilde{\omega}^P_{ROW}(x)$ and $\widetilde{\omega}^P_F(x)$, whereas the actual external tariff t affects the investment and sourcing decisions. We call the term in brackets the *relocation function*:

$$r(x, t) \equiv \Psi_t(\widetilde{\omega}^P_F(x), t) - \Psi_0(\widetilde{\omega}^P_{ROW}(x)).$$

The relocation function captures the change in welfare due to a buyer who, induced by a tariff preference of size x, abandons a vertical chain with supplier $\widetilde{\omega}^P_{ROW}(x)$ in ROW and forms a new one with supplier $\widetilde{\omega}^P_F(x)$ in *Foreign*, but invests and produces according to the external tariff t. In turn, the function $\phi(x; \gamma_F, G)$ captures the flow rate of buyers that (due to the PTA) move from vertical chains with ROW suppliers with productivity $\widetilde{\omega}^P_{ROW}(x)$ to new vertical chains with *Foreign* suppliers with productivity $\widetilde{\omega}^P_F(x)$. Specifically, we have

$$\phi(x; \gamma_F, G) \equiv \frac{\gamma_F (1 - \gamma_F) g(\widetilde{\omega}^P_F(x)) g(\widetilde{\omega}^P_{ROW}(x))}{\gamma_F g(\widetilde{\omega}^P_F(x)) + (1 - \gamma_F) g(\widetilde{\omega}^P_{ROW}(x))}.$$

The flow rate is the product of the densities of the ROW and *Foreign* cutoff suppliers, divided by the weighted average of the two densities. The extensive-margin welfare effect $XM(0, \gamma_F)$ aggregates the relocation function over all supplier relocations that occur under the PTA according to the weights $\phi(x; \gamma_F, G)$.

The following lemma characterizes $r(x, t)$.

Lemma 5 *Let $\gamma_H = 0$. The relocation function $r(x, t)$ has the following properties:*
(i) $r(0, t) = \Delta\Psi(\widetilde{\omega}_\beta, t)$.
(ii) $r(x, t)$ is strictly decreasing in x.
(iii) $r(t, t) = -t q^*_P \left(\widetilde{\omega}^P_F(t)\right) < 0$.

The relocation function is defined such that $r(0, t) = \Delta\Psi(\widetilde{\omega}_\beta, t)$. This is the welfare effect for the cutoff supplier under no PTA; hence, $r(0, t)$ essentially gives the welfare effect at the boundary

[26] See the Appendix for the derivation of this expression.

between the intensive and extensive margins. This may be positive or negative. Then, $r(x, t)$ falls as x increases. As x increases, the productivity of the old ROW supplier $\widetilde{\omega}^P_{ROW}(x)$ improves and the productivity of the new *Foreign* supplier $\widetilde{\omega}^P_F(x)$ worsens. Hence, the productivity gap between old and new suppliers grows with x. This lowers the welfare effect of relocation for two reasons: directly, as a lower-productivity supplier generates less social surplus under any given trade regime; and indirectly, because we know from Lemma 3 that the relationship-strengthening effects of a PTA is weaker for lower-productivity suppliers.

We also have that, at $x = t$, $r(t, t)$ is unambiguously negative. The net joint profits are, under the PTA, the same for supplier $\widetilde{\omega}^P_F(t)$ in *Foreign* and supplier $\widetilde{\omega}^P_{ROW}(t)$ in ROW. Since the difference between social welfare and joint profits is tariff revenue (which unambiguously falls with the PTA), $r(t, t)$ represents lost tariff revenue under the PTA, evaluated for the least productive new *Foreign* supplier that forms a vertical chain: $-tq^*_t\left(\widetilde{\omega}^P_F(t)\right)$. Hence, the matching process induces welfare losses for sure at the margin, even after accounting for potentially beneficial changes in investment.

Now consider the effect of the tariff. If $r(0, t) = \Delta\Psi(\widetilde{\omega}_\beta, t)$ is negative, then $XM(0, \gamma_F) < 0$ regardless of how specialized suppliers are distributed.

Proposition 9 *In a North-South PTA with $\gamma_H = 0$, if $t \geq 2\widehat{t}(\widetilde{\omega}_\beta)$, the extensive-margin welfare effect is negative for any distribution of specialized suppliers.*

If $t < 2\widehat{t}(\widetilde{\omega}_\beta)$, then the welfare analysis is more complicated. For this case, Figure 7 illustrates the intensive-margin welfare effect, the extensive-margin welfare effect, and their relationships to matching diversion and overall welfare. When we make these comparisons, it is helpful to change variables in the r function once more. We can write

$$XM(0, \gamma_F) = \gamma_F \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}^P_F(t)} r_A(\omega, t)dG(\omega),$$

where

$$r_A(\omega, t) \equiv \Psi_t(\omega, t) - \Psi_0(\widetilde{\omega}^P_{ROW}(\omega)).$$

The *adjusted relocation function* $r_A(\omega, t)$ shows, for an arbitrary external tariff t, the welfare impact of the PTA due to each relocation to ω in *Foreign* from $\widetilde{\omega}^P_{ROW}(\omega)$ in ROW. For $\omega \leq \widetilde{\omega}_\beta$, $\Delta\Psi(\omega, t)$ denotes the impact due to each incumbent supplier in *Foreign*. The $r_A(\omega, t)$ function is decreasing in ω, and $r_A(\widetilde{\omega}_\beta, t) = \Delta\Psi(\widetilde{\omega}_\beta, t)$ yields the boundary between the intensive-margin welfare effect (aggregating over $\Delta\Psi(\omega, t)$ from 0 to $\widetilde{\omega}_\beta$) and the extensive-margin welfare effect (aggregating over $r_A(\omega, t)$ from $\widetilde{\omega}_\beta$ to $\widetilde{\omega}^P_F(t)$).

In the figure, $t < 2\widehat{t}(\widetilde{\omega}_\beta)$, so that $r_A(\widetilde{\omega}_\beta, t) = \Delta\Psi(\widetilde{\omega}_\beta, t) > 0$. In this case, the extensive-margin welfare effect *could* be positive.[27] But of course we know that it includes the negative matching-diversion effect. To see the effects separately, note that the dashed line is the welfare impact that the PTA would have for suppliers distributed over $(\widetilde{\omega}_\beta, \widetilde{\omega}^P_F(t)]$ *if they were incumbent*. But they are

[27]If $t \geq 2\widehat{t}(\widetilde{\omega}_\beta)$, the kink in the function would occur at or below the x-axis, $\Delta\Psi(\widetilde{\omega}_\beta, t) = r_A(\widetilde{\omega}_\beta, t) \leq 0$, and $XM(0, \gamma_F)$ would be negative for sure (Proposition 9).

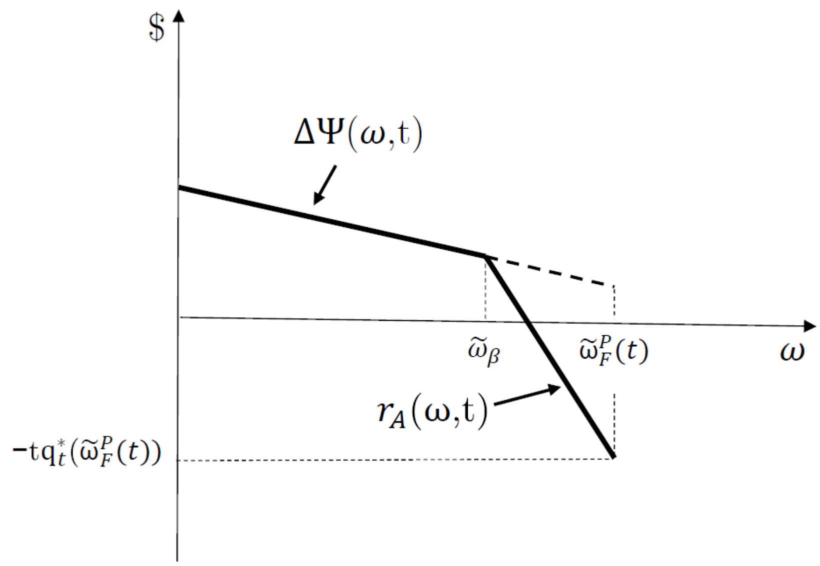

Fig. 7: Welfare Effects and the Adjusted Relocation Function

Notes: The figure shows the welfare effect of the PTA due to a single Y-chain in *Foreign*, conditional on the inverse productivity (ω) of the supplier under the PTA. The external tariff satisfies $t < 2t(\widetilde{\omega}_\beta)$. For $\omega \in [0, \widetilde{\omega}_\beta]$, the supplier is active both with and without the PTA, and the welfare effect is $\Delta\Psi(\omega, t)$. For $\omega \in (\widetilde{\omega}_\beta, \widetilde{\omega}_F^P(t)]$, the supplier is active only under the PTA, replaces a higher-productivity supplier from ROW, and the welfare effect is $r_A(\omega, t)$.

not. Instead, they replace suppliers distributed over $[\widetilde{\omega}_{ROW}^P(t), \widetilde{\omega}_\beta)$ previously matched in ROW. The difference between the dashed line and the solid line to the right of $\widetilde{\omega}_\beta$ represents exactly the loss due to matching diversion. This effect is small for the very first rematches, but grows large as relocation continues. As t rises, the $r_A(\omega, t)$ portion of the curve necessarily lengthens, since $\widetilde{\omega}_F^P$ increases with t.

Ultimately, $XM(0, \gamma_F)$ is negative whenever the matching-diversion effect dominates. This depends upon the density. And as we show in the Appendix, for certain densities, the matching-diversion effect dominates for *any* t. Intuitively, this holds for any density where the flow rate of suppliers at the margin weakly increases as the tariff increases. An example is the uniform density.

Nevertheless, it is always possible [if $t < 2\widehat{t}(\widetilde{\omega}_\beta)$] to construct a density such that the mass of suppliers near $\widetilde{\omega}_\beta$ is sufficiently higher than the mass near $\widetilde{\omega}_F^P(t)$, so that $XM(0, \gamma_F)$ is positive. We have the following example.

Example 2 *Let* $t < 2\widehat{t}(\widetilde{\omega}_\beta)$, $\gamma_H = 0, \gamma_F = \frac{1}{2}$ *and*

$$g_{PU}(\omega) = \begin{cases} \frac{1-2\widehat{\varepsilon}\eta}{1-2\widehat{\varepsilon}} & \text{if} \quad \omega \in [0, \beta p_w - \widehat{\varepsilon}) \\ \eta & \text{if} \quad \omega \in [\beta p_w - \widehat{\varepsilon}, \beta p_w + \widehat{\varepsilon}] \\ \frac{1-2\widehat{\varepsilon}\eta}{1-2\widehat{\varepsilon}} & \text{if} \quad \omega \in (\beta p_w + \widehat{\varepsilon}, p_w] \end{cases},$$

where $\eta \in \left(0, \frac{1}{2\widehat{\varepsilon}}\right)$ *and*

$$\widehat{\varepsilon} = \frac{t\left[2p_w(1-\beta)\alpha(1-\alpha)b^2 - t\left[2c - 2\alpha b^2 + \alpha^2 b^2\right]\right]}{\left[4p_w(1-\beta)(2c-\alpha^2 b^2) + 2t\alpha(1-\alpha)b^2\right]} > 0.$$

This distribution is piecewise uniform, with three different regions. Equilibrium matching yields $\widetilde{\omega}^P_{ROW}(t) = \beta p_w - \frac{t}{2}$ *in the low-*ω *region of* $g(\omega)$, $\widetilde{\omega}_\beta = \beta p_w$ *in the center of the middle-*ω *region, and* $\widetilde{\omega}^P_F(t) = \beta p_w + \frac{t}{2}$ *in the high-*ω *region. This specification is constructed specifically so that* $r_A(\widetilde{\omega}_\beta + \widehat{\varepsilon}, t) = 0$*. Then*

$$XM\left(0, \frac{1}{2}\right) = \frac{1}{2}\left[\eta \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_\beta + \widehat{\varepsilon}} r_A(\omega, t)d\omega + \left(\frac{1 - 2\widehat{\varepsilon}\eta}{1 - 2\widehat{\varepsilon}}\right) \int_{\widetilde{\omega}_\beta + \widehat{\varepsilon}}^{\widetilde{\omega}^P_F} r_A(\omega, t)d\omega\right].$$

It follows that $\int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_\beta+\widehat{\varepsilon}} r_A(\omega,t)d\omega > 0$ *and* $\int_{\widetilde{\omega}_\beta+\widehat{\varepsilon}}^{\widetilde{\omega}^P_F} r_A(\omega,t)d\omega < 0$*. Hence, for* η *sufficiently close to* $\frac{1}{2\widehat{\varepsilon}}$*, the extensive-margin effect is positive.*

Figure 8 highlights the intuition. If the density of idle suppliers (under no PTA) in Foreign *is very high for supplier relocations very close to* $\widetilde{\omega}_\beta$*, and this density is very low for other supplier relocations, then it is possible to have a positive extensive-margin effect even when* $\gamma_H = 0$*. Compare Figure 8 with a uniform density* $g_{k1}(\omega) = \frac{1}{p_w}$*. The density* $g_{PU}(\omega)$ *distorts* $g_{k1}(\omega)$*, allocating more density near* $\widetilde{\omega}_\beta$ *and less density near* $\widetilde{\omega}^P_{ROW}$ *and* $\widetilde{\omega}^P_F$*. But it does not alter the equilibrium cutoffs* $\widetilde{\omega}^P_{ROW}, \widetilde{\omega}_\beta$ *and* $\widetilde{\omega}^P_F$*. This reflects a situation where (1)* Foreign *has a large number of suppliers with productivity near* $\widetilde{\omega}_\beta$ *that are idle without the PTA, but relatively few less-productive idle suppliers; and (2) most* ROW *suppliers that are replaced also have productivity near* $\widetilde{\omega}_\beta$*.*

Note that in this example, if $t > 2\widehat{t}(\widetilde{\omega}_\beta)$*, then no positive* $\widehat{\varepsilon}$ *exists and it is impossible to construct a density that yields* $XM(0, \gamma_F) > 0$*.*

The full welfare effect of the North-South case also includes $IM(0, \gamma_F)$, whose sign depends on the balance between the relationship-strengthening and the sourcing-diversion effects over all existing vertical chains in *Foreign*, as discussed in subsection 5.2. The analysis of this term follows that of the "large natural partner" ($\gamma_F = 1$) case studied in the previous subsection, which is also a special North-South case. Typically, a North-South PTA raises aggregate welfare when intensive-margin welfare effects are sufficiently strong relative to any negative extensive-margin welfare effects. While the net effect of those forces is generally ambiguous, some forces tilt the balance in one direction or the other. In the Appendix, we show that the change in welfare is positive for sufficiently low t, is maximized for a level of t strictly below $\widehat{t}_W^{\gamma_F=1}$, and is negative for sufficiently high t. This is similar to the "large natural partner" case. The key difference is that because of extensive-margin effects, $\Delta W(0, \gamma_F)$ may not be globally concave.

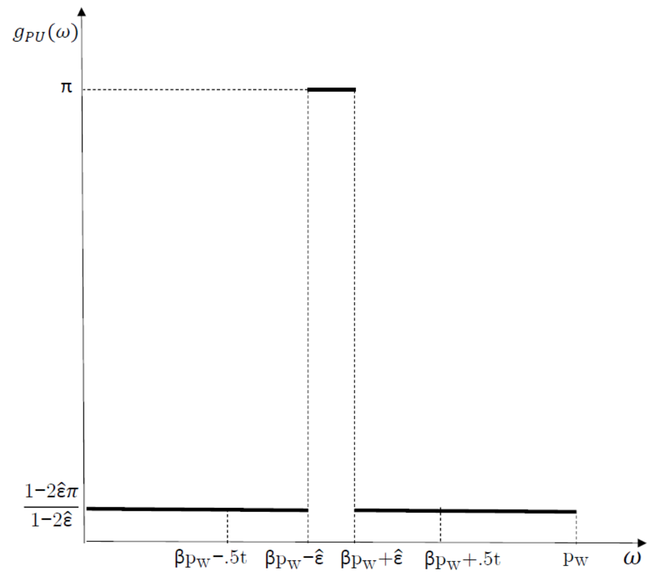

Fig. 8: A Distribution that Yields a Positive Extensive-Margin Effect in the North-South Case

Notes: This figure illustrates how the extensive-margin welfare effect may be positive even without matching creation. For η sufficiently close to $\frac{1}{2\widehat{\varepsilon}}$, the mass of suppliers is concentrated very close to the no PTA cutoff supplier, $\widetilde{\omega}_{ROW}^{N} = \widetilde{\omega}_{F}^{N} = \beta p_w$. Hence, most "new" suppliers in *Foreign* under the PTA (and most "old" suppliers in ROW) are very close to $\omega = \beta p_w$. If $t < 2\widehat{t}(\beta p_w)$, the welfare effect for the $\omega = \beta p_w$ supplier is positive, $r_A(\beta p_w, t) = \Delta\Psi(\beta p_w, t) > 0$, and the full welfare effect $XM(0, \gamma_F)$ is likewise positive.

7 Alternative Model Specifications and Empirical Implications

7.1 Single Sourcing and Within-Chain Trade Creation

Throughout the analysis, we assume that the range of parameters is such that all buyers engage in dual sourcing. That assumption keeps Q^* fixed, shutting down within-chain trade creation in intermediates (and, in turn, insulating the market for final goods from the PTA). Accordingly, dual sourcing within Y-chains isolates the within-chain tradeoff between the relationship-strengthening and the sourcing-diversion effects. If that restriction were relaxed, then a buyer might purchase all Q^* inputs in equilibrium from the specialized supplier. This would occur in the vertical chains with the most-productive specialized suppliers, whose marginal cost functions are sufficiently low to intersect $V'(Q)$ below p_w.

Under single sourcing, the welfare effects of the PTA due to an incumbent vertical chain in *Foreign* are different. The supplier's marginal cost curve falls due to the removal of t from the input price. The buyer purchases more inputs, conditional on a given investment. Because there is no dual sourcing, the equilibrium number of inputs is efficient and there is no sourcing-diversion effect.[28] The supplier's investment increases with the PTA but is below the first-best—the PTA

[28] The bargaining surplus is the same as faced by a social planner, minus a constant term that reflects the buyer's

will never cause over-investment under single sourcing. Hence, in vertical chains where the supplier productivity is high enough for given demand $V'(Q)$ and tariff t, in the sense that the marginal cost curve intersects $V'(Q)$ below p_w, there is no within-chain tradeoff between the relationship-strengthening and the sourcing-diversion effects. Instead, there is just a pure trade-creation (or "sourcing-creation") effect within incumbent vertical chains in *Foreign*, driven by the reduced marginal cost of inputs. Furthermore, because Q^* rises, the production of the differentiated final good increases, raising consumer surplus, just as standard Vinerian trade creation would.

Generally, within an industry we have three qualitatively different cases for the welfare effects of a PTA due to a single vertical chain. The chains with the most productive suppliers yield pure sourcing creation and provide a positive contribution to welfare. Those with the least productive suppliers correspond to the case we fully analyze in the paper. As we have seen, *among them*, the most productive suppliers tend to yield a positive contribution to welfare, but the least productive suppliers tend to lower the welfare impact of the PTA. There is also a third group of suppliers, with intermediate levels of productivity, who operate under dual sourcing in the absence of a PTA but under single sourcing under a PTA. They generate correspondingly intermediate welfare impacts.

Hence, in the general case, the intensive-margin welfare implications of a PTA would be analogous to those in our previous analysis. We would have again that the most productive vertical chains tend to generate a higher welfare gain under the PTA, except that now that rationale is reinforced by the gains coming from the pure-sourcing-creation chains (and potentially also by gains from the intermediary group).

Putting all together, then, from a qualitative perspective we lose little by focusing on the dual-sourcing case, while stressing the within-chain tradeoff between the relationship-strengthening and the sourcing-diversion effects. In a possible quantitative analysis of our mechanisms, on the other hand, it would be necessary to expand the analysis to the more general case.

7.2 Alternative Patterns of Trade for the Generics Industry

As indicated in section 2, the assumption that all generic inputs are produced in ROW is not without loss of generality. Here we briefly discuss how our results would be affected under alternative specifications for the location of the generic industry.

The structure that would introduce the greatest changes to our results is when *Foreign* is an exporter of z. In that case, the reduction of tariffs with the PTA would affect both types of inputs in the same way, and therefore would have no impact on the sourcing of q from *Foreign*. However, an analogous, but in some aspects inverse, analysis could be made for the sourcing of q from ROW, which would then be discriminated against z under the PTA. There would be, in particular, a relationship-*weakening* effect for vertical chains preserved in ROW after the PTA.

If instead *Home* exported z, its domestic industry would provide all generic inputs that its buyers purchase. The price of z in *Home* would then be p_w with and without a PTA. In that case, production and investment decisions of *Home* specialized suppliers would be insensitive to

option to purchase only generic inputs if bargaining breaks down.

the tariff, a feature not particularly appealing in our context. To offer effective protection to *Home* suppliers in such a setting, we would then need to allow *Home* to treat its specialized suppliers more favorably than its generic suppliers, e.g. by subsidizing specialized input production. If a PTA were to extend this subsidy to *Foreign* suppliers, then a subsidy of magnitude t would have the same effects on production and investment decisions that a tariff of magnitude t has in the current version of the model.

With both of these alternative structures, the more general points of our analysis would nevertheless remain valid. First, a PTA improves the incentives to invest for the specialized suppliers whose inputs become cheaper than the generic alternative, but worsens the incentives to invest for the specialized suppliers whose inputs become relatively more expensive than the generic alternative because of the tariff preference. Second, since for the relocation of vertical chains what matters is the relative tariff on specialized inputs across locations, the essential insights from our current analysis would remain entirely unchanged, Indeed, it is worth stressing that all the important action in the model hinges upon the *difference* between the tariffs applied to z and q. For notational simplicity, we defined the initial tariff to be the same, t, for both, but the analysis would carry through if the tariffs on z and q were, respectively, t_z and t_q, with $t_z \neq t_q$.

There are other possible specifications for the role of the generic industry, but they would have even less impact on our results. For example, consider *Home* were an importer but produced some z domestically. Since the price of z at *Home* would remain $p_w + t$ with and without a PTA, the welfare analysis would not be altered in any important way. Another possibility is when *Foreign* has an industry of generics but the industry is unable to supply enough z to fulfill *Home*'s demand, so *Home* still imports z from *ROW* under the PTA. Again, that would leave all of our novel results essentially unchanged, for the reasons discussed above. The main difference is that in this case the PTA would also generate standard trade diversion in the sourcing of generic inputs, of the type analyzed by Grossman and Helpman (1995).

7.3 Positive Implications of a PTA

As discussed in the introduction, our model is designed to capture some key features of the phenomenon often described as "global sourcing." That is, its building blocs are defined precisely to reflect the empirical regularities present in this new, increasingly relevant environment. Now, are the *implications* of the model also empirically relevant?

Measuring the welfare consequences of trade agreements is notoriously difficult. Welfare is not directly measurable and trade agreements (and the external tariffs in PTAs) are chosen by governments, and therefore endogenous. Hence, for such a task, quantitative models are needed. While challenging, adaptation of existing frameworks (e.g., a la Caliendo and Parro, 2015) to the main features of our model seem feasible.

Nevertheless, our model also has some clear positive, testable implications for the matching structure of the economy, for the productivity of matched firms, and for the trade flows following the formation of a PTA. The effects depend on the location of the match prior to the PTA.

Specifically, the model implies that buyers forming vertical chains in PTA member countries prior to the agreement keep their original suppliers and source more from them. Moreover, because of the higher investment levels, the observed productivity of those incumbent suppliers should increase. Also, the increase in investment is especially large for the new suppliers inside the bloc, who did not invest and did not export before the PTA. Therefore, a key testable prediction of the model is that there should be a particularly large increase in investment for average-productivity producers that start to export because of preferential market access. This is similar to what Lileeva and Trefler (2010) find in the context of preferential liberalization between Canada and the United States.

Now, for firms forming vertical chains in non-PTA countries and in the PTA importing country (*Home*) prior to the PTA, there will not be any change for those buying from the highly productive suppliers there. In turn, those sourcing from less productive firms outside the PTA switch to suppliers within the trading bloc, and their baseline productivity is lower than the productivity of their previous suppliers outside the bloc. The same happens for partnerships in *Home*.

We could summarize those predictions as follows. Upon the formation of a PTA:

(a) Incumbent suppliers in the PTA exporting country should increase investment, export more, and display higher productivity than they did before the agreement.

(b) The increase in investment should be particularly high for the new suppliers in the PTA exporting country, who did not export before the agreement.

(c) There should be no effect, including in trade flows, for high-productivity incumbent suppliers in the PTA importing country and in those outside the bloc.

(d) Pre-PTA relationships with the lowest-productivity suppliers in the PTA importing country and outside the bloc should be replaced by new ones in the PTA exporting country.

This set of testable implications makes the model falsifiable. Recently, datasets that include the identity and characteristics of matched firms across countries are becoming increasingly available. If a PTA is implemented between two countries for which such data are available, one could readily investigate the validity of those implications. Naturally, one would need to identify cases where one can deal with the endogeneity of the agreement and of its external and preferential tariffs (which may depend on whether the agreement is reciprocal or formed under the Enabling Clause), and would also need to control for issues like firm organization (i.e., whether the firm is vertically integrated), which we sidestep in our theoretical analysis. Lileeva and Trefler (2010) offer a useful approach to estimate investment effects upon the formation of a PTA.

Sugita, Teshima and Seira (2018) provide an interesting analysis along those lines, focused on the characteristics of the matching equilibria. They study the effects of a trade policy shock that is akin to a removal of import preferences: the end of very restrictive import quotas on (some) clothing and textiles products on 1 January 2005 in the US. Those quotas applied to imports coming from some countries (especially China) but not to others (like Mexico). The authors investigate how the trade policy shock affected the structure of buyer-seller matches between the US and Mexico. They find that the removal of the preferential treatment that Mexico enjoyed caused significant partner switching, and that those changes played the main role in the ensuing trade flow adjustments. Interestingly, they also find that the trade shock increased the efficiency of the matches. In the context of our paper, one could interpret their results as evidence that there was matching diversion under the preferential quota system, which receded once the quotas were eliminated.

8 Conclusion

Our goal is to provide a framework for studying the implications of preferential liberalization in the context of global sourcing, where contracts are incomplete, buyers and suppliers match and trade customized inputs, and investments are relationship-specific. We uncover several forces that standard welfare analyses of PTAs, by not considering those elements, are likely to miss. In particular, we show that a PTA affects the efficiency of the production process both through cost-reducing investment and through changes in the set of active vertical chains. Some of those changes enhance global welfare, while others have the opposite effect. We characterize those forces in detail and indicate the conditions that make one or the other prevail.

An important element behind the design of our framework is tractability. This makes it possible to extend the model in several directions. For example, one could readily accommodate alternative patterns of organization, such as vertical integration. It is well known that firms may integrate to overcome hold-up problems (Grossman and Hart, 1986), and that discriminatory protection may alter those incentives (Ornelas and Turner, 2008). In our basic framework, with all else equal, integrated firms would tend to overinvest under discriminatory protection and more productive firms would have stronger incentives to integrate. Incorporating integration would enrich the overall analysis of the welfare effects of PTAs, but would also introduce an extensive taxonomy of cases. We therefore leave it for future research, noting that the primary challenge will be to capture the key sources of heterogeneity determining integration (e.g., productivity, integration/governance costs, etc.). Another, empirically relevant, type of heterogeneity that would be interesting to study in future work is buyer heterogeneity, which would require a more elaborated matching structure, e.g. including assortative matching.

Our framework could be employed to shed light on current policy debates as well. For example, in the recent renegotiation of NAFTA, its members agreed to tighten the rules of origin (ROO) requirements for the automotive sector to qualify for zero tariffs within the bloc. As Conconi et al. (2018) show, NAFTA's existing ROOs already reduce imports of intermediate products from outside the bloc. Here, one way to incorporate the new tightening would be to explicitly model the sourcing of additional inputs by specialized suppliers. ROOs would bind when the privately optimal volume of specialized inputs under the PTA is below the minimum requirement level of input imports to make the final good qualify for free trade within the PTA—naturally, this would require explicit modeling of the market of final goods. The firms' tradeoff would be whether to distort their trade of customized inputs, so that the final good enjoys duty-free access in the market of the PTA partner.

At a more general level, an increasingly important theme for policymakers and academics alike is the expansion of global value chains. Our results help to justify the view that PTAs promote the intensification of GVCs. First, they generate "more depth" in existing relationships, fueled by more investment. Second, PTAs also generate "more width," in the sense of fueling the formation of new relationships. Now, our setting is very simple, with a supply chain containing just two specialized firms plus a competitive fringe. In contrast, a typical GVC includes several producers and parts cross several national borders. But as Yi (2003) points out, tariffs are typically applied on gross exports. This suggests that the mechanisms we develop are likely to be even more important for "genuine" GVCs, like the ones studied by Antràs and de Gortari (2020).

Baldwin (2011), the World Trade Organization (2011) and several others have argued that regionalism nowadays is about the rules that underpin fragmentation of production, not about preferential market access. As such, Baldwin (2011) claims that the traditional Vinerian approach is outdated and that we need "a new framework that is as simple and compelling as the old one, but relevant to 21st century regionalism" (p. 23). Here we introduce several features that are deemed central for the international fragmentation of production, and yet show that preferential market access remains key for understanding the welfare impact of PTAs—probably more than it has ever been for the trade of final goods. Intuitively, deep provisions in PTAs interact with preferential market access in many ways. We are optimistic that future extensions of our model can be used to study the welfare implications of such provisions as well. Hence, one could view our model as a first step towards a framework that extends the Vinerian view to the "new regionalism" world.

References

Antràs, Pol (2020). "Conceptual Aspects of Global Value Chains." *World Bank Economic Review*, forthcoming.

Antràs, Pol and Davin Chor (2013). "Organizing the global value chain." *Econometrica* 81, 2127-2204.

Antràs, Pol and Alonso de Gortari (2020). "On the Geography of Global Value Chains." *Econometrica*, forthcoming.

Antràs, Pol, Teresa Fort and Felix Tintelnot (2017)."The Margins of Global Sourcing: Theory and Evidence From U.S. Firms." *American Economic Review* 107(9), 2514-2564.

Antràs, Pol and Elhanan Helpman (2004). "Global sourcing." *Journal of Political Economy* 112, 552-580.

Antràs, Pol and Elhanan Helpman (2008). "Contractual frictions and global sourcing." In E. Helpman, D. Marin, and T. Verdier (eds.), *The Organisation of Firms in a Global Economy*, 9-54, Boston: Harvard University Press.

Antràs, Pol and Robert Staiger (2012a). “Offshoring and the Role of Trade Agreements.” *American Economic Review* 102(7), 3140-3183.

Antràs, Pol and Robert Staiger (2012b). “Offshoring and the Role of Trade Agreements.” *American Economic Review Papers & Proceedings* 102(3), 470-476.

Bagwell, Kyle, Chad Bown and Robert Staiger (2016). “Is the WTO Passé?” *Journal of Economic Literature* 54(4), 1125-1231.

Baier, Scott and Jeffrey Bergstrand (2004). “Economic determinants of free trade agreements.” *Journal of International Economics* 64, 29-63.

Baier, Scott, Jeffrey Bergstrand and Michael Feng (2014). “Economic integration agreements and the margins of international trade.” *Journal of International Economics* 93, 339-350.

Baldwin, Richard (2011). "21st century regionalism: filling the gap between 21st century trade and 20th century trade rules." World Trade Organization, Staff Working Paper ERSD-2011-08.

Baldwin, Richard (2016). "The World Trade Organization and the Future of Multilateralism." *Journal of Economic Perspectives* 30(1), 95-116.

Bernard, Andrew, Andreas Moxnes and Karen Helene Ullveit-Moe (2018). “Two-sided Heterogeneity and Trade.” *Review of Economics and Statistics* 100, 424-439.

Bhagwati, Jagdish and Arvind Panagariya (1996). “Preferential trading areas and multilateralism: strangers, friends or foes?” In J. Bhagwati and A. Panagariya (eds.), *The Economics of Preferential Trade Agreements*, 1-78. Washington, DC: AEI.

Blanchard, Emily (2015). “A Shifting Mandate: International Ownership, Global Fragmentation, and a Case for Deeper Integration under the WTO.” *World Trade Review* 14(1), 87-99.

Blanchard, Emily, Chad Bown and R. Johnson (2017). “Global Supply Chains and Trade Policy." Mimeo.

Blaum, Joaquin, Michael Peters and Claire Lelarge (2018). “The Gains from Input Trade with Heterogeneous Importers.” *American Economic Journal: Macroeconomics* 10, 77-127.

Boehm, Johannes and Ezra Oberfield (2020). “Misallocation in the Market for Inputs: Enforcement and the Organization of Production.” *Quarterly Journal of Economics*, forthcoming.

Bustos, Paula (2011). “Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms.” *American Economic Review* 101(1), 304-340.

Caliendo, Lorenzo and Fernando Parro (2015). “Estimates of the Trade and Welfare Effects of NAFTA.” *The Review of Economic Studies* 82(1), 1-44.

Conconi, Paola, Manuel García-Santana, Laura Puccio and Roberto Venturini (2018). “From Final Goods to Inputs: The Protectionist Effect of Rules of Origin.” *American Economic Review* 108(8), 2335-2365.

Freund, Caroline and Emanuel Ornelas (2010). "Regional Trade Agreements." *Annual Review of Economics* 2, 139-166.

Gretsky, Neil, Joseph Ostroy and William Zame (1992). "The nonatomic assignment model." *Economic Theory* 2(1), 103-127.

Grossman, Sanford and Oliver Hart (1986). "The costs and benefits of ownership: a theory of vertical and lateral integration." *Journal of Political Economy* 94, 691-719.

Grossman, Gene and Elhanan Helpman (1995). "The politics of free-trade agreements." *American Economic Review* 85, 667-90.

Grossman, Gene and Elhanan Helpman (2005). "Outsourcing in a global economy." *Review of Economic Studies* 72, 135-159.

Grossman, Gene and Elhanan Helpman (2020). "When Tariffs Disrupt Global Supply Chains." Mimeo.

Heise, Sebastian, Justin Pierce, Georg Schaur and Peter K. Schott (2017). "Trade Policy and the Structure of Supply Chains." Mimeo.

Johnson, Robert and Guillermo Noguera (2017). "A Portrait of Trade in Value Added Over Four Decades." *The Review of Economics and Statistics* 99(5), 896-911.

Lileeva, Alla and Daniel Trefler (2010). "Improved Access to *Foreign* Markets Raises Plant-Level Productivity ... for Some Plants." *Quarterly Journal of Economics* 125(3), 1051-1099.

Limao, Nuno (2016). "Preferential Trade Agreements." In: K. Bagwell & R. W. Staiger (eds.), *Handbook of Commercial Policy*, Elsevier/North Holland, Volume 1B, Chapter 6, 279-367.

Maggi, Giovanni (2014). "International Trade Agreements." In: G. Gopinath, E. Helpman and K. Rogoff (eds.), *Handbook of International Economics*, Elsevier/North Holland, Volume 4, Chapter 6, 317-390.

McLaren, John (2000). "Globalization and Vertical Structure." *American Economic Review* 90, 1239-1254.

Ornelas, Emanuel and John Turner (2008). "Trade liberalization, outsourcing and the hold-up problem." *Journal of International Economics* 74, 225-241.

Ornelas, Emanuel and John Turner (2012). "Protection and international sourcing." *Economic Journal* 122: 26-63.

Qiu, Larry and Barbara Spencer (2002). "Keiretsu and relationship-specific investments: implications for market-opening policy." *Journal of International Economics* 58, 49-79.

Ruta, Michele (2017). "Preferential trade agreements and global value chains: Theory, evidence, and open questions." In: Global Value Chain Development Report, *Measuring and Analyzing the Impact of GVCs on Economic Development*, 175-185. World Bank, Washington DC.

Sugita, Yoichi, Kensuke Teshima and Enrique Seira (2018). “Assortative Matching of Exporters and Importers.” Mimeo.

Viner, Jacob (1950). *The Customs Union Issue.* New York: Carnegie Endowment for International Peace.

World Trade Organization (2011). *Preferential Trade Agreements and the WTO: from Co-existence to Coherence.* World Trade Report 2011.

Yi, Kei-Mu (2003). “Can vertical specialization explain the growth of world trade?” *Journal of Political Economy* 111(1), 52-102.

Appendix

I Proofs

Proof of Lemma 1. Using (5) and (9), we have that

$$HUP_0 = i^e - i_0^* = \frac{2bc\,(1-\alpha)\,(p_w-\omega)}{(2c-b^2)\,(2c-\alpha b^2)},$$

which is clearly decreasing in ω. ■

Proof of Lemma 2. Using (15) and (9), we have that

$$EXC_t = i_t^* - i^e = -\frac{2bc\,(1-\alpha)\,(p_w-\omega)}{(2c-b^2)\,(2c-\alpha b^2)} + \Delta i,$$

which is clearly increasing in ω. ■

Proof of Lemma 3. Substituting expressions (10) and (18) into (37) and differentiating yields $\frac{\partial \Delta\Psi_{RS}}{\partial\omega} = \frac{-2b^2\alpha(1-\alpha)t}{(2c-\alpha b^2)^2}$. Differentiating (35) yields $\frac{\partial\Delta\Psi(\omega,t)}{d\omega} = \frac{-2b^2\alpha(1-\alpha)t}{(2c-\alpha b^2)^2}$. Hence, $\frac{\partial\Delta\Psi_{RS}}{\partial\omega} = \frac{\partial\Delta\Psi(\omega,t)}{d\omega} < 0$. ■

Proof of Lemma 4. Taking the first-order condition for the relationship-strengthening effect, we have

$$\frac{d\Delta\Psi_{RS}}{dt} = \left(\frac{2c-b^2}{c}\right)\frac{\alpha b^2}{(2c-\alpha b^2)^2}\left[\frac{2(1-\alpha)c}{(2c-b^2)}(p_w-\omega) - \alpha t\right] = 0.$$

This is clearly positive at $t=0$. Solving it yields

$$\widehat{t}^R(\omega) = \frac{2(1-\alpha)(p_w-\omega)c}{\alpha\,(2c-b^2)} > 0.$$

The second derivative is

$$\frac{d^2\Delta\Psi_{RS}}{dt^2} = -\left(\frac{2c-b^2}{c}\right)\frac{\alpha^2 b^2}{(2c-\alpha b^2)^2} < 0.$$

Hence, $\Delta\Psi_{RS}$ is strictly concave, so $\widehat{t}^R(\omega)$ is a unique maximum and $\Delta\Psi_{RS}$ is an inverted-U function of t. The sourcing-diversion effect, $\Delta\Psi_{SD} = -\frac{t^2}{2c}$, is clearly strictly decreasing in t.

Now consider the overall within-chain effect, $\Delta\Psi(\omega, t)$. Differentiating (35) with respect to t, we find

$$\frac{d\Delta\Psi(\omega,t)}{dt} = \frac{2}{(2c-\alpha b^2)^2}\left[b^2\alpha(1-\alpha)(p_w-\omega) - t(2c+\alpha^2b^2-2\alpha b^2)\right]. \tag{44}$$

This is clearly positive for $t=0$, so $\Delta\Psi(\omega,t)>0$ for sufficiently low t. Solving $\frac{d\Delta\Psi(\omega,t)}{dt}=0$ yields $\widehat{t}(\omega)$. The second -order condition is

$$\frac{d^2\Delta\Psi(\omega,t)}{dt^2} = -\left[\frac{2(2c-2\alpha b^2+\alpha^2b^2)}{(2c-\alpha b^2)^2}\right] < 0.$$

Hence, $\Delta\Psi(\omega,t)$ is strictly concave and achieves a unique maximum at $\widehat{t}_\omega$. This shows that $\Delta\Psi(\omega,t)$ is an inverted-U function of t.

Finally, note that $t = 2\widehat{t}(\omega)$ solves $\Delta\Psi(\omega,t)=0$. Then it follows immediately that $\Delta\Psi(\omega,t)>0$ if $t \in \left(0, 2\widehat{t}(\omega)\right)$ and $\Delta\Psi(\omega,t)<0$ if $t > 2\widehat{t}(\omega)$. ■

Proof of Proposition 2. It is clear from (40) that $AWC > 0$ if $\Delta\Psi(\omega,t) \geq 0$ for all ω. A sufficient condition for this is that $\Delta\Psi(\widetilde{\omega}_F^{MID}, t) \geq 0$. Generally, $\widetilde{\omega}_F^{MID}$ depends upon distribution G, but $\widetilde{\omega}_F^{MID} \leq \widetilde{\omega}_F^P \leq \widetilde{\omega}_\beta + t$. Thus, a sufficient condition for a positive AWC is that $\Delta\Psi(\widetilde{\omega}_\beta + t, t) \geq 0$. Using (35), we can write

$$\Delta\Psi(\widetilde{\omega}_\beta + t, t) = \frac{t}{(2c-\alpha b^2)^2}\left[2b^2\alpha(1-\alpha)(p_w-\widetilde{\omega}_\beta - t) - t(2c+\alpha^2b^2-2\alpha b^2)\right]. \tag{45}$$

The positive tariff that makes this expression equal zero is

$$t_{AWC} = \frac{2b^2\alpha(1-\alpha)(p_w-\widetilde{\omega}_\beta)}{2c-\alpha^2b^2}.$$

The term in brackets in (45) is clearly decreasing in t. Hence, $\Delta\Psi(\widetilde{\omega}_F^{MID}, t) \geq 0$ for any $t \leq t_{AWC}$, and so is AWC. Finally, if $t \geq 2\widehat{t}(0)$, then by Lemma 4 we have that $\Delta\Psi(\omega,t) < 0$ for any ω and $AWC < 0$, completing the proof. ■

Proof of Proposition 4. Because $\Psi_0(\omega)$ is strictly decreasing in ω and G is continuous and strictly increasing in ω, it follows that

$$\begin{aligned}
\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} \Psi_0(\omega)dG(\omega) &< \left[\gamma_F \int_{\widetilde{\omega}_{ROW}^N}^{\widetilde{\omega}_F^{MID}} dG(\omega)\right]\Psi_0(\widetilde{\omega}_{ROW}^N) \\
&= \left[(1-\gamma_H-\gamma_F)\int_{\widetilde{\omega}_{ROW}^P}^{\widetilde{\omega}_{ROW}^N} dG(\omega)\right]\Psi_0(\widetilde{\omega}_{ROW}^N) \\
&< (1-\gamma_H-\gamma_F)\int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_0(\omega,t)dG(\omega),
\end{aligned}$$

where the equality uses the definition of $\widetilde{\omega}_F^{MID}$ from equation (33). This completes the proof. ■

Proof of Proposition 5. Clearly, the overall welfare effect of a PTA is zero for $t = 0$. To show it is positive for low t, we will show that this effect is increasing in t in the limit as $t \to 0$.

First, let us calculate a few important derivatives for future reference. Noting that the natural cutoff supplier $\widetilde{\omega}_\beta$ obtains in all three countries when $t = 0$, we can solve to find

$$\frac{d\widetilde{\omega}_H^N}{dt} \mid_{t=0} = \frac{(1-\gamma_H)g(\widetilde{\omega}_H^N - t)}{\gamma_H g(\widetilde{\omega}_H^N) + (1-\gamma_H)g(\widetilde{\omega}_H^N - t)} \mid_{t=0} = (1-\gamma_H) > 0;$$

$$\frac{d\widetilde{\omega}_F^N}{dt} \mid_{t=0} = \frac{d\widetilde{\omega}_{ROW}^N}{dt} = \frac{-\gamma_H g(\widetilde{\omega}_F^N + t)}{\gamma_H g(\widetilde{\omega}_F^N + t) + (1-\gamma_H)g(\widetilde{\omega}_F^N)} \mid_{t=0} = -\gamma_H < 0;$$

$$\frac{d\widetilde{\omega}_H^P}{dt} \mid_{t=0} = \frac{d\widetilde{\omega}_F^P(t)}{dt} = \frac{(1-\gamma_H-\gamma_F)g(\widetilde{\omega}_F^P - t)}{(\gamma_H+\gamma_F)g(\widetilde{\omega}_F^P) + (1-\gamma_H-\gamma_F)g(\widetilde{\omega}_F^P - t)} \mid_{t=0} = (1-\gamma_H-\gamma_F) > 0;$$

$$\frac{d\widetilde{\omega}_{ROW}^P}{dt} \mid_{t=0} = \frac{-(\gamma_H+\gamma_F)g(\widetilde{\omega}_{ROW}^P + t)}{(\gamma_H+\gamma_F)g(\widetilde{\omega}_{ROW}^P + t) + (1-\gamma_H-\gamma_F)g(\widetilde{\omega}_{ROW}^P)} \mid_{t=0} = -(\gamma_H+\gamma_F) < 0.$$

Now, using the expression for $\Delta W(\gamma_H, \gamma_F)$ in equation (31), we find that the derivative of the intensive-margin effect is

$$\frac{dIM}{dt}|_{t=0} = \gamma_F \left[\int_0^{\widetilde{\omega}_F^N(t)} \frac{d\Delta\Psi(\omega,t)}{dt} dG(\omega) + \Delta\Psi(\widetilde{\omega}_F^N, t) g(\widetilde{\omega}_F^N) \frac{d\widetilde{\omega}_F^N}{dt} \right] |_{t=0} = \gamma_F \int_0^{\widetilde{\omega}_\beta} \frac{d\Delta\Psi(\omega,0)}{dt} dG(\omega) > 0.$$

where the latter follows from the fact that $\Delta\Psi(\widetilde{\omega}_\beta, 0) = 0$.

In turn, the derivative of the first term inside the brackets in equation (31) is

$$\frac{d}{dt}\left(\gamma_F \int_{\widetilde{\omega}_F^N}^{\widetilde{\omega}_F^P} \Psi_t(\omega,t) dG(\omega)\right) = \gamma_F \left[\int_{\widetilde{\omega}_F^N}^{\widetilde{\omega}_F^P} \frac{dt\Psi_t(\omega,t)}{dt} dG(\omega) \right.$$
$$\left. + \Psi_t(\widetilde{\omega}_F^P, t) g(\widetilde{\omega}_\beta) \frac{d\widetilde{\omega}_F^P}{dt} - \Psi_t(\widetilde{\omega}_F^N, t) g(\widetilde{\omega}_\beta) \frac{d\widetilde{\omega}_F^N}{dt} \right].$$

Evaluating at $t = 0$, it becomes

$$\frac{d}{dt}(\bullet)|_{t=0} = \gamma_F(1-\gamma_F)\Psi_t(\widetilde{\omega}_\beta, 0) g(\widetilde{\omega}_\beta) > 0.$$

The derivative of the second term inside the brackets is

$$\frac{d}{dt}\left(-\gamma_H \int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \Psi_t(\omega,t) dG(\omega)\right) = -\gamma_H \left[\int_{\widetilde{\omega}_H^P}^{\widetilde{\omega}_H^N} \frac{d\Psi_t(\omega,t)}{dt}|_{t=0} dG(\omega) \right.$$
$$\left. + \Psi_t(\widetilde{\omega}_H^N, t) g(\widetilde{\omega}_H^N) \frac{d\widetilde{\omega}_H^N}{dt} - \Psi_t(\widetilde{\omega}_H^P, t) g(\widetilde{\omega}_H^P) \frac{d\widetilde{\omega}_H^P}{dt} \right].$$

Evaluating at $t=0$, it becomes

$$\frac{d}{dt}(\bullet)\,|_{t=0} = -\gamma_H\gamma_F\Psi_t(\widetilde{\omega}_\beta,0)g(\widetilde{\omega}_\beta) < 0.$$

Finally, the derivative of the third term inside the brackets is

$$\frac{d}{dt}\left(-(1-\gamma_H-\gamma_F)\int_{\widetilde{\omega}^P_{ROW}}^{\widetilde{\omega}^N_{ROW}}\Psi_0(\omega)dG(\omega)\right) = -(1-\gamma_H-\gamma_F)\left[\int_{\widetilde{\omega}^P_{ROW}}^{\widetilde{\omega}^N_{ROW}}\frac{d\Psi_0(\omega)}{dt}dG(\omega)\right.$$
$$\left.+\Psi_0(\widetilde{\omega}^N_{ROW})g(\widetilde{\omega}^N_{ROW})\frac{d\widetilde{\omega}^N_{ROW}}{dt} - \Psi_0(\widetilde{\omega}^P_{ROW})g(\widetilde{\omega}^P_{ROW})\frac{d\widetilde{\omega}^P_{ROW}}{dt}\right].$$

Evaluating at $t=0$, it becomes

$$\frac{d}{dt}(\bullet)\,|_{t=0} = -(1-\gamma_H-\gamma_F)\gamma_F\Psi_0(\widetilde{\omega}_\beta)g(\widetilde{\omega}_\beta)$$

Collecting terms, we have

$$\begin{aligned}
\frac{d\Delta W}{dt}|_{t=0} &= \gamma_F\int_0^{\widetilde{\omega}_\beta}\frac{d\Delta\Psi(\omega,0)}{dt}dG(\omega) + \gamma_F(1-\gamma_F)\Psi_t(\widetilde{\omega}_\beta,0)g(\widetilde{\omega}_\beta) \\
&\quad -\gamma_H\gamma_F\Psi_t(\widetilde{\omega}_\beta,0)g(\widetilde{\omega}_\beta) - (1-\gamma_H-\gamma_F)\gamma_F\Psi_0(\widetilde{\omega}_\beta)g(\widetilde{\omega}_\beta) \\
&= \gamma_F\int_0^{\widetilde{\omega}_\beta}\frac{d\Delta\Psi(\omega,0)}{dt}dG(\omega) + \gamma_F(1-\gamma_H-\gamma_F)\left[\Psi_t(\widetilde{\omega}_\beta,0) - \Psi_0(\widetilde{\omega}_\beta)\right]g(\widetilde{\omega}_\beta) \\
&= \gamma_F\int_0^{\widetilde{\omega}_\beta}\frac{d\Delta\Psi(\omega,0)}{dt}dG(\omega) > 0.
\end{aligned}$$

Hence, in the limit as $t \to 0$, the change in welfare with respect to t is positive; the overall welfare effect of a PTA is therefore positive for sufficiently low t. ■

Proof of Proposition 6. Let $t \leq 2t(\widetilde{\omega}_\beta)$. Then by Lemma 4, $\Delta\Psi(\widetilde{\omega}_\beta,t) \geq 0$. From (32), we know that $\widetilde{\omega}_F^{MID} = \widetilde{\omega}_F^N \leq \widetilde{\omega}_\beta$. Then Lemma 3 implies that $\Delta\Psi(\widetilde{\omega}_F^N,t) \geq \Delta\Psi(\widetilde{\omega}_\beta,t) \geq 0$, and that $AWC(\gamma_H, 1-\gamma_H) = (1-\gamma_H)\int_0^{\widetilde{\omega}_F^N}\Delta\Psi(\omega,t)dG(\omega) \geq 0$. ■

Proof of Proposition 7. Let $\gamma_F = 1$. By definition, the welfare impact of the PTA is zero when $t=0$. When there is a small increase in t, $AWC(0,1)$ changes according to $\frac{\partial AWC(0,1)}{\partial t} = \int_0^{\widetilde{\omega}_\beta}\frac{\partial\Delta\Psi(\omega,t)}{\partial t}dG(\omega)$. We have from (44) that $\frac{\partial\Delta\Psi(\omega,t)}{\partial t} = \frac{2}{(2c-\alpha b^2)^2}\left\{-t\left[2c-2\alpha b^2+\alpha^2b^2\right] + (p_w-\omega)\alpha(1-\alpha)b^2\right\}$. This expression is strictly positive when evaluated at $t=0$. Therefore, for sufficiently low preference margins, $AWC(0,1) > 0$. Now notice that $\frac{\partial^2 AWC(0,1)}{\partial t^2} = \int_0^{\widetilde{\omega}_\beta}\frac{\partial^2\Delta\Psi(\omega,t)}{\partial t^2}dG(\omega) = -\int_0^{\widetilde{\omega}_\beta}\frac{2\left[2c-2\alpha b^2+\alpha^2b^2\right]}{(2c-\alpha b^2)^2}dG(\omega) < 0$. Therefore, $AWC(0,1)$ is maximized when $\frac{\partial AWC(0,1)}{\partial t} = 0$. Simple algebra shows that this happens when $t = \hat{t}_W^{\gamma_F=1}$. Finally, after some manipulation it follows that, when $t = 2\hat{t}_W^{\gamma_F=1}$, $AWC(0,1) = 0$. Since $\frac{\partial^2 AWC(0,1)}{\partial t^2} < 0$, $AWC(0,1) < 0$ when $t > 2\hat{t}_W^{\gamma_F=1}$. Because $0 < \mathbf{E}(\omega;\omega \leq \widetilde{\omega}_\beta) < \widetilde{\omega}_\beta$, it also follows that $\hat{t}(\widetilde{\omega}_\beta) < \hat{t}_W^{\gamma_F=1} < \hat{t}(0)$. ■

Proof of Proposition 8. Equilibrium matching when $\gamma_F = 1$ requires $G_1(\widetilde{\omega}_{\beta 1}) = \beta$ and

$G_2(\widetilde{\omega}_{\beta 2}) = \beta$. If $G_2(\omega)$ **FOSD** $G_1(\omega)$, the two distributions satisfy $G_1(\omega) \geq G_2(\omega)$. It follows that $\widetilde{\omega}_{\beta 1} \leq \widetilde{\omega}_{\beta 2}$. The changes in welfare from the PTA for the two distributions are

$$\begin{aligned} \Delta W_1(0,1;G_1) &= \int_0^{\widetilde{\omega}_{\beta 1}} \Delta\Psi(\omega,t)dG_1(\omega) \text{ and} \\ \Delta W_2(0,1;G_2) &= \int_0^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega,t)dG_2(\omega). \end{aligned}$$

Hence,

$$\Delta\Delta W \equiv \Delta W_1(0,1;G_1) - \Delta W_2(0,1;G_2) = \int_0^{\widetilde{\omega}_{\beta 1}} \Delta\Psi(\omega,t)dG_1(\omega) - \int_0^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega,t)dG_2(\omega).$$

Integrating both terms by parts, we can write

$$\begin{aligned} \Delta\Delta W &= \Delta\Psi(\omega,t)G_1(\omega)|_0^{\widetilde{\omega}_{\beta 1}} - \int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_1(\omega)d\omega - \left[\Delta\Psi(\omega,t)G_2(\omega)|_0^{\widetilde{\omega}_{\beta 2}} - \int_0^{\widetilde{\omega}_{\beta 2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_2(\omega)d\omega\right] \\ &= \Delta\Psi(\widetilde{\omega}_{\beta 1},t)G_1(\widetilde{\omega}_{\beta 1}) - \int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_1(\omega)d\omega - \left[\Delta\Psi(\widetilde{\omega}_{\beta 2},t)G_2(\widetilde{\omega}_{N2}) - \int_0^{\widetilde{\omega}_{\beta 2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_2(\omega)d\omega\right] \\ &= \beta\left[\Delta\Psi(\widetilde{\omega}_{\beta 1},t) - \Delta\Psi(\widetilde{\omega}_{\beta 2},t)\right] - \int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[G_1(\omega) - G_2(\omega)]d\omega + \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_2(\omega)d\omega \\ &= \left\{\beta\left[\Delta\Psi(\widetilde{\omega}_{\beta 1},t) - \Delta\Psi(\widetilde{\omega}_{\beta 2},t)\right] + \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}G_2(\omega)d\omega\right\} - \int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[G_1(\omega) - G_2(\omega)]d\omega. \end{aligned}$$

Because $\frac{d\Delta\Psi(\omega,t)}{d\omega} < 0$, it follows that

$$-\int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[G_1(\omega) - G_2(\omega)]d\omega > 0.$$

Hence, it remains to show that the term in curly brackets is positive. Integrating its second term by parts, we can write

$$\begin{aligned} \{\cdot\} &= \beta\left[\Delta\Psi(\widetilde{\omega}_{\beta 1},t) - \Delta\Psi(\widetilde{\omega}_{\beta 2},t)\right] + \Delta\Psi(\omega,t)G_2(\omega)|_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} - \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega,t)dG_2(\omega) \\ &= \beta\left[\Delta\Psi(\widetilde{\omega}_{\beta 1},t) - \Delta\Psi(\widetilde{\omega}_{\beta 2},t)\right] + \Delta\Psi(\widetilde{\omega}_{\beta 2},t)G_2(\widetilde{\omega}_{\beta 2}) - \Delta\Psi(\widetilde{\omega}_{\beta 1},t)G_2(\widetilde{\omega}_{\beta 1}) - \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega,t)dG_2(\omega) \\ &= \beta\Delta\Psi(\widetilde{\omega}_{\beta 1},t) - \Delta\Psi(\widetilde{\omega}_{\beta 1},t)G_2(\widetilde{\omega}_{\beta 1}) - \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega,t)dG_2(\omega), \end{aligned}$$

where the final line comes from setting $G_2(\widetilde{\omega}_{\beta 2}) = \beta$ and simplifying. We then have

$$\begin{aligned} \{\cdot\} &= \Delta\Psi(\widetilde{\omega}_{\beta 1}, t)\left[G_2(\widetilde{\omega}_{\beta 2}) - G_2(\widetilde{\omega}_{\beta 1})\right] - \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \Delta\Psi(\omega, t) dG_2(\omega) \\ &= \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \left[\Delta\Psi(\widetilde{\omega}_{\beta 1}, t) - \Delta\Psi(\omega, t)\right] dG_2(\omega) > 0. \end{aligned}$$

Hence,

$$\Delta\Delta W = \int_{\widetilde{\omega}_{\beta 1}}^{\widetilde{\omega}_{\beta 2}} \left[\Delta\Psi(\widetilde{\omega}_{\beta 1}, t) - \Delta\Psi(\omega, t)\right] dG_2(\omega) - \int_0^{\widetilde{\omega}_{\beta 1}} \frac{d\Delta\Psi(\omega, t)}{d\omega}[G_1(\omega) - G_2(\omega)]d\omega > 0,$$

concluding the proof. ■

Proof of Lemma 5. For part (i), note that $\widetilde{\omega}_F^P(0) = \widetilde{\omega}_{ROW}^P(0) = \widetilde{\omega}_\beta$, so $r(0, t) = \Psi_t(\widetilde{\omega}_\beta, t) - \Psi_0(\widetilde{\omega}_\beta) = \Delta\Psi(\widetilde{\omega}_\beta, t)$. For part (ii), differentiate to find

$$\frac{dr(x, t)}{dx} = \frac{d\Psi_t(\widetilde{\omega}_F^P(x), t)}{d\widetilde{\omega}_F^P} \frac{d\widetilde{\omega}_F^P}{dx} - \frac{\Psi_0(\widetilde{\omega}_{ROW}^P(x))}{d\widetilde{\omega}_{ROW}^P} \frac{d\widetilde{\omega}_{ROW}^P}{dx},$$

which is negative because $\frac{d\Psi_t(\widetilde{\omega}_F^P(x),t)}{d\widetilde{\omega}_F^P} < 0$, $\frac{d\widetilde{\omega}_F}{dx} > 0$, $\frac{\Psi_0(\widetilde{\omega}_{ROW}^P(x))}{d\widetilde{\omega}_{ROW}^P} < 0$ and $\frac{d\widetilde{\omega}_{ROW}^P}{dx} < 0$. For part (iii), note that $r(t, t) = \Psi_t(\widetilde{\omega}_F^P(t), t) - \Psi_0(\widetilde{\omega}_{ROW}^P(t)) = \Psi_t(\widetilde{\omega}_F^P(t), t) - \Psi_0(\widetilde{\omega}_F^P(t) - t)$. Simple algebra shows that this equals $-tq_t^*\left(\widetilde{\omega}_F^P(t)\right)$. ■

Proof of Proposition 9. Suppose $t > 2\widehat{t}(\widetilde{\omega}_\beta) = \frac{2\alpha(1-\alpha)b^2\left[p_w - \widetilde{\omega}_\beta\right]}{2c - 2\alpha b^2 + \alpha^2 b^2}$. Then

$$\Delta\Psi(\widetilde{\omega}_\beta, t) = \frac{t}{(2c - \alpha b^2)^2}\left[2b^2\alpha(1-\alpha)(p_w(1-\beta)) - t(2c + \alpha^2 b^2 - 2\alpha b^2)\right] < 0.$$

By Lemma 3, it follows that $\int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_F^P(t)} \Delta\Psi(\omega, t) dG(\omega) < 0$. Therefore, $XM(0, \gamma_F) < 0$. ■

II Efficient investment levels

Without an agreement, the efficient investment level solves

$$\max_i p_w q_0 - C(q_0, i, \omega) - I(i). \tag{46}$$

The first-order necessary condition is

$$p_w \frac{dq_0}{di} - C_q(q_0, i, \omega)\frac{dq_0}{di} - C_i(q_0, i, \omega) = I'(i).$$

Using (3), this expression simplifies to $-C_i(q_0, i, \omega) = I'(i^e)$, as indicated in (8).

With a PTA, the efficient investment level also solves (46), after replacing q_0 with q_t. The first-order necessary condition is analogous to the one above, but simplifies to

$$-t\frac{dq_t}{di} - C_i(q_t, i, \omega) = I'(i).$$

This expression may appear to yield a level of investment different from i^e. However, developing it further we obtain

$$-t\frac{b}{c} + b\left(\frac{p_w + t - \omega + bi}{c}\right) = 2i,$$

which is satisfied exactly when $i = i^e$.

III Explicit expressions for welfare

Inserting equilibrium investments and levels of inputs, we have the following expressions for the welfare generated by a single Y-chain without and with preferential treatment:

$$\begin{aligned}\Psi_0(\omega) &= [V(Q^*) - p_w Q^*] + \frac{(p_w - \omega)^2 \left(2c - \alpha^2 b^2\right)}{\left(2c - \alpha b^2\right)^2}, \\ \Psi_t(\omega, t) &= [V(Q^*) - p_w Q^*] + \frac{(p_w + t - \omega)^2 \left(2c - \alpha^2 b^2\right)}{\left(2c - \alpha b^2\right)^2} - \frac{2t\,(p_w + t - \omega)}{(2c - \alpha b^2)}.\end{aligned}$$

Observe that the term in brackets does not change with the level of discriminatory protection.

IV The matching equilibrium

We describe the full details of a Walrasian equilibrium in the market for matches. Equilibrium requires an assignment of buyers to suppliers and a fee schedule describing the net transfer from each supplier to the buyer that she is matched to, such that buyers and suppliers choose matches to maximize profits (taking the schedule as given) and the market for matches clears. For both the no-PTA and PTA cases, we first introduce a more general notation and state equilibrium conditions using this notation, then convert back to the notation in the main text. Let $l \in \{F, H, ROW\}$ denote the country in which a supplier is located. Let the level of discriminatory protection for a B-S pair be $\tau \in \{0, t\}$.

No PTA Under no PTA (N), discriminatory protection maps one-to-one with the supplier's location l in the following way: for $l = H$, we have $\tau = t$; for $l \in \{F, ROW\}$, we have $\tau = 0$. Let the suppliers pay the buyers a matching fee $M : \Omega \times [0, \beta] \times \{0, t\} \to \mathbb{R}$. Let the assignment of matches under no PTA follow $\mu_N : [0, \beta] \to \Omega \times \{F, H, ROW\}$. Define the gross utility for a buyer of type $\zeta \in [0, \beta]$ matched with a supplier of type ω, where the match enjoys discriminatory protection via tariff $\tau \in \{0, t\}$, as $U^B(\zeta, \omega, \tau)$. Define the gross utility for a supplier of type ω matched with a buyer of type ζ, where the match enjoys discriminatory protection via tariff $\tau \in \{0, t\}$, as $U^S(\omega, \tau, \zeta)$. Three sets of conditions must hold:

1. For each buyer $\zeta \in [0, \beta]$, the assignment $\mu_N(\zeta)$ solves

$$\max_{\{\omega,\tau\}} U^B(\zeta, \omega, \tau) + M(\omega, \zeta, \tau)$$

Given the fee, buyers maximize profits over a choice of supplier (productivity ω and discriminatory protection τ).

2. For each supplier $(\omega, \tau) \in \Omega \times \{0, t\}$, each buyer match $\zeta \in \mu_N^{-1}(\Omega, \{F, H, ROW\})$ solves

$$i) \ \max_{\{\zeta\}} U^S(\omega, \tau, \zeta) - M(\omega, \zeta, \tau).$$

Given the fee, suppliers maximize profits over a choice of buyer. Because there is an excess of suppliers, there is an additional requirement:

$$ii) \ \max_{\{\zeta\}} U^S(\omega, \tau, \zeta) - M(\omega, \zeta, \tau) \leq 0 \ if \ \mu_N^{-1}(\Omega, \{F, H, ROW\}) \text{ is empty.}$$

If a supplier is unassigned, then her payoff from matching with a buyer would be non-positive.

3. The assignments must also match all available buyers to all suppliers with types more productive than marginal types:

$$\begin{aligned} \int_{\mu_N([0,\beta]\times\{H,F,ROW\})} dG(\omega) &= \beta, \\ \int_{\mu_N([0,\beta]\times H)} dG(\omega) &\leq \gamma_H \\ \int_{\mu_N([0,\beta]\times F)} dG(\omega) &\leq \gamma_F, \\ \int_{\mu_N([0,\beta]\times ROW)} dG(\omega) &\leq 1 - \gamma_H - \gamma_F. \end{aligned}$$

Statement 1 implies that $\frac{dM(\omega,\zeta,\tau)}{d\omega} = \frac{-dU^B(\zeta,\omega,\tau)}{d\omega}$. Hence, $\frac{dM(\omega,\zeta,0)}{d\omega} = \frac{-dU^B(\zeta,\omega,0)}{d\omega}$ and $\frac{dM(\omega,\zeta,t)}{d\omega} = \frac{-dU^B(\zeta,\omega,t)}{d\omega}$. Because U^S is a constant function of ζ, statement 2(i) implies that $M(\omega, \zeta, \tau)$ is a constant function of ζ. Statement 2(ii) implies the marginal supplier earns exactly zero profit. Hence, we drop the ζ arguments from all functions and drop the tariff argument from functions when $\tau = 0$. We write gross utilities under $\tau = 0$ as $U_0^S(\omega, 0) \equiv U_0^S(\omega, 0, \zeta)$ and $U_0^B(\omega) \equiv U_S(\zeta, \omega, 0)$. For the $\tau = t$ case, we write gross utilities as functions of the tariff size, $U_t^S(\omega, t) \equiv U^S(\omega, t, \zeta)$ and $U_t^B(\omega, t) \equiv U^B(\zeta, \omega, t)$. For the fees, we write $M_l^N(\omega) \equiv M(\omega, \zeta, 0)$ for $l \in \{F, ROW\}$ and $M_H^N(\omega, t) = M(\omega, \zeta, t)$ for $l = H$.

Denoting $\widetilde{\omega}_y^N$ as the marginal supplier in country l under no PTA, we have $U_0^S(\widetilde{\omega}_l^N) = M_l^N(\widetilde{\omega}_l^N)$ for $l \in \{F, ROW\}$ and $U_t^S(\widetilde{\omega}_H^N, t) = M_H^N(\widetilde{\omega}_H^N, t)$ for $l = H$. The former condition implies $\widetilde{\omega}_F^N = \widetilde{\omega}_{ROW}^N$. We also know from statement 1 that $U_t^B(\widetilde{\omega}_H^N, t) + M_H^N(\widetilde{\omega}_H^N, t) = U_0^B(\widetilde{\omega}_l^N) + M_l^N(\widetilde{\omega}_l^N)$ for $l \in \{F, ROW\}$, because otherwise some buyers would not be maximizing profits. Hence, substituting, we can write, for $l \in \{F, ROW\}$, $U_0^B(\widetilde{\omega}_l^N) + U_0^S(\widetilde{\omega}_l^N) = U_t^B(\widetilde{\omega}_H^N, t) + U_t^S(\widetilde{\omega}_H^N, t)$. It then follows immediately from the multiplicative separability of the total profit from a $B-S$ pair that $\widetilde{\omega}_F^N = \widetilde{\omega}_{ROW}^N = \widetilde{\omega}_H^N - t$.[29] Statement 3 implies that $\gamma_H G(\widetilde{\omega}_H^N) + (1-\gamma_H) G(\widetilde{\omega}_H^N - t) = \beta$.

For $l \in \{F, ROW\}$, because $\frac{dM_l^N(\omega)}{d\omega} = \frac{-dU_0^B(\omega)}{d\omega}$, we also have that $M_l^N(\omega)$ may be written as the sum of $-U_0^B(\omega)$ and a term that does not vary with ω. Similarly, $M_H^N(\omega, t)$ may be written as the sum of $U_t^B(\omega, t)$ and a term that does not vary with ω. We construct the fees by specifying, for $l \in \{F, ROW\}$, $M_l^N(\omega) = -U_0^B(\omega) + \varphi_{N,l}$ and $M_H^N(\omega, t) = -U_t^B(\omega, t) + \varphi_{N,H}$. Returning to the marginal suppliers, we can then write, for $l \in \{F, ROW\}$, $U_0^S(\widetilde{\omega}_l^N) = -U_0^B(\widetilde{\omega}_l^N) + \varphi_{N,l}$ and $U_t^S(\widetilde{\omega}_H^N, t) = -U_t^B(\widetilde{\omega}_H^N, t) + \varphi_{N,H}$. Hence, we can solve for $\varphi_{N,H}$ and substitute to find

$$M_H^N(\omega, t) = U_t^S(\widetilde{\omega}_H^N, t) - \left[U_t^B(\omega, t) - U_t^B(\widetilde{\omega}_H^N, t)\right].$$

For $l \in \{F, ROW\}$, we can similarly solve for $k_{N,l}$ and substitute to find

$$M_l^N(\omega, t) = U_0^S(\widetilde{\omega}_l^N) - \left[U_0^B(\omega) - U_0^B(\widetilde{\omega}_l^N)\right].$$

PTA Under a PTA (P), the level of discriminatory protection for a B-S pair, $\tau \in \{0, t\}$, maps one-to-one with the supplier's location l in the following way: for $l \in \{F, H\}$, we have $\tau = t$; for $l = ROW$, we have $\tau = 0$. Conditions 1-3 must again hold for a matching assignment $\mu_P : [0, \beta] \to \Omega \times \{F, H, ROW\}$ that differs from μ_N because suppliers in country F enjoy discriminatory protection. Denote fees under the PTA as $M_l^P(\omega, t) = M(\omega, \zeta, t)$ for $l \in \{F, H\}$ and $M_{ROW}^P(\omega) \equiv M(\omega, \zeta, 0)$. Similarly, denote $\widetilde{\omega}_l^P$ as the marginal supplier in country l under a PTA. Applying conditions 1-3 using arguments analogous to those above, we find that $\widetilde{\omega}_F^P = \widetilde{\omega}_H^P = \widetilde{\omega}_{ROW}^P + t$ and that these cutoffs must also satisfy $(\gamma_H + \gamma_F) G(\widetilde{\omega}_H^P) + (1 - \gamma_H - \gamma_F) G(\widetilde{\omega}_H^P - t) = \beta$. Constructing constant terms for the fees analogously, we also find, for $l \in \{F, H\}$,

$$M_l^P(\omega, t) = U_t^S(\widetilde{\omega}_l^P, t) - \left[U_t^B(\omega, t) - U_t^B(\widetilde{\omega}_l^P, t)\right],$$

and for $l = ROW$,

$$M_{ROW}^P(\omega, t) = U_0^S(\widetilde{\omega}_{ROW}^P) - \left[U_0^B(\omega) - U_0^B(\widetilde{\omega}_{ROW}^P)\right].$$

Stability Our setting is a continuous assignment model. Hence, equilibrium yields a stable matching (Gretsky, Ostroy and Zame, 1992).

[29] This separability is net of the buyer's default payoff $[V(Q^*) - (p_w + t)Q^*]$, which is a constant term. The separability is easily seen by plugging into (20) and (19) for both $\tau = 0$ and $\tau = t$, and adding the expressions together.

V Rewriting $XM(0, \gamma_F)$ using a change of variables

Start with the expression for the extensive-margin effect:

$$XM(0, \gamma_F) \equiv \gamma_F \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_F^P(t)} \Psi_t(\omega, t) g(\omega) d\omega - (1 - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P(t)}^{\widetilde{\omega}_\beta} \Psi_0(\omega) g(\omega) d\omega.$$

Changing the variable from ω to x, note that $d\omega = d\widetilde{\omega}_F^P(x)dx$, so that

$$dx = \frac{d\omega}{d\omega_F(x)}.$$

Then note that

$$\gamma_F g(\widetilde{\omega}_F^P(x)) d\omega = \frac{\phi(x; \gamma_F, G) d\omega}{d\widetilde{\omega}_F^P(x)} = \phi(x; \gamma_F, G) dx,$$

where the first equality follows from

$$d\widetilde{\omega}_F^P(x) = \frac{(1 - \gamma_F) g(\widetilde{\omega}_{ROW}^P(x))}{\gamma_F g(\widetilde{\omega}_F^P(x)) + (1 - \gamma_F) g(\widetilde{\omega}_{ROW}^P(x))}.$$

Substituting back in and adjusting the bounds of integration ($\widetilde{\omega}_\beta$ to $x = 0$ at the lower end and $\widetilde{\omega}_F^P$ to $x = t$ at the upper end), we then have that

$$\gamma \int_{\widetilde{\omega}_\beta}^{\widetilde{\omega}_F^P} \Psi_t(\omega, t) g(\omega) d\omega = \int_0^t \Psi_t(\widetilde{\omega}_F^P(x), t) \phi(x; \gamma_F, G) dx.$$

A similar manipulation of the second term in $XM(0, \gamma_F)$ yields

$$(1 - \gamma_F) \int_{\widetilde{\omega}_{ROW}^P(t)}^{\widetilde{\omega}_\beta} \Psi_0(\omega) g(\omega) d\omega = \int_0^t \Psi_0(\widetilde{\omega}_{ROW}^P(x)) \phi(x; \gamma_F, G) dx.$$

Hence, we obtain expression (43) from the main text:

$$XM(0, \gamma_F) = \int_0^t \left[\Psi_t(\widetilde{\omega}_F^P(x), t) - \Psi_0(\widetilde{\omega}_{ROW}^P(x)) \right] \phi(x; \gamma_F, G) dx.$$

VI The Welfare Impact of a North-South PTA

The analysis of the full welfare impact includes the intensive and extensive margins. When $t < 2\widehat{t}(\widetilde{\omega}_\beta)$, the welfare effect of the PTA is positive for *all* intensive-margin suppliers, which means that the effect is also positive for some part of the extensive margin (recall Figure 7). However, even in this case, there are conditions where the total extensive-margin welfare effect is always negative. We now state a monotonicity condition.

Condition 1 *The flow rate $\phi(x; \gamma_F, F)$ is weakly increasing in x.*

Condition 1 implies that, as the tariff increases, the flow rate of relocations (weakly) increases. For a continuously differentiable density, this is equivalent to assuming that

$$(1-\gamma_F)g(\widetilde{\omega}^P_{ROW}(x))^3 g'(\widetilde{\omega}^P_F(x)) - \gamma_F g(\widetilde{\omega}^P_F(x))^3 g'(\widetilde{\omega}^P_{ROW}(x)) \geq 0.$$

With a uniform distribution, $g_{k1}(\omega) = \frac{1}{p_w}$, the flow rate of new relocations is constant and satisfies Condition 1 for any γ_F and t. For other distributions, such as $g_{k2}(\omega) = \frac{2\omega}{p_w^2}$, it is often the case that it holds for some γ_F and t, but not all. Still, if Condition 1 holds, then $XM(0,\gamma_F) < 0$ regardless of t.

Proposition 10 *Under Condition 1, the extensive-margin welfare effect for a North-South PTA is negative for any positive t.*

Proof. We use $XM(0,\gamma_F) = \int_0^t r(x,t)\phi(x;\gamma_F,G)dx$. It is obvious that if $t = 0$, then $XM(0,\gamma_F) = 0$. Differentiating, we have

$$\frac{dXM(0,\gamma_F)}{dt} = r(t,t)\phi(t;\gamma_F,G) + \int_0^t \frac{dr(x,t)}{dt}\phi(x;\gamma_F,G)dx.$$

Because $r(0,0) = 0$, it is obvious that $\frac{dXM(0,\gamma_F;t=0)}{dt} = 0$. Then, if $\frac{d^2XM(0,\gamma_F)}{dt^2} < 0$ for all t, $XM(0,\gamma_F)$ is strictly concave and therefore negative for all t as well. We now show that, under Condition 1, $\frac{d^2XM(0,\gamma_F)}{dt^2} < 0$ for all t. After using the functional form for the r function to substitute, we have

$$\begin{aligned} \frac{d^2XM(0,\gamma_F)}{dt^2} &= \left\{\phi(t;\gamma_F,G)\left[\left(\frac{d}{dt}\frac{-2t\left(p_w - \widetilde{\omega}^P_{ROW}(t)\right)}{(2c-\alpha b^2)}\right) + \frac{d\Psi_t(t,\omega)}{dt}\right]\right\} \\ &\quad + \int_0^t \frac{d^2r(x,t)}{dt^2}\phi(\gamma_F,x;G)dx - \left[\frac{2t\left(p_w - \widetilde{\omega}^P_{ROW}(t)\right)}{(2c-\alpha b^2)}\left(\frac{d\phi(t;\gamma_F,F)}{dt}\right)\right]. \end{aligned} \tag{47}$$

Start with the term in braces, expand the expression and substitute according to the functional form for $\frac{d\Psi_t(t,\omega)}{dt}$:

$$\begin{aligned} \{\cdot\} &= \left(\frac{-2\phi(t;\gamma_F,G)}{(2c-\alpha b^2)}\right) \\ &\quad \left\{\left[t\left(-\frac{d\widetilde{\omega}^P_{ROW}(t)}{dt}\right) + \left(p_w - \widetilde{\omega}^P_{ROW}(t)\right)\right] - \left[\frac{\alpha(1-\alpha)b^2\left(p_w - \widetilde{\omega}^P_F(t)\right) - t\left(2c - 2\alpha b^2 + \alpha^2 b^2\right)}{2c-\alpha b^2}\right]\right\}. \end{aligned}$$

Rearranging, we can write

$$\{\cdot\} = \left(\frac{-2\phi(t;\gamma_F,G)}{(2c-\alpha b^2)}\right)$$
$$\left\{\left[\left(p_w-\widetilde{\omega}^P_{ROW}(t)\right)-\left(\frac{\alpha(1-\alpha)b^2\left(p_w-\widetilde{\omega}^P_F(t)\right)}{2c-\alpha b^2}\right)\right]+t\left[\left(-\frac{d\widetilde{\omega}^P_{ROW}(t)}{dt}\right)+\frac{(2c-2\alpha b^2+\alpha^2b^2)}{2c-\alpha b^2}\right]\right\}.$$

The term in the second bracket $[\cdot]$ is clearly positive, and a few lines of algebra show that the term in the first bracket is also positive. Hence, the entire expression is negative.

Next consider the first term on the second line of (47). This is the aggregate of the second-order effects of the tariff for relocations, each of which is negative. Hence, $\int_0^t \frac{d^2r(x,t)}{dt^2}\phi(\gamma_F,x;G)dx<0$.

Finally, consider the second term on the second line of (47). Because $\frac{d\phi(t;\gamma_F,G)}{dt}\geq 0$ under Condition 1, the entire term is non-positive. This shows that $\frac{d^2XM(0,\gamma_F)}{dt^2}<0$. ■

Intuitively, if the flow rate of relocations rises with t, then there are relatively more relocations at the margin than inframarginally. As a result, any welfare improvement from higher investments is dominated by welfare losses due to matching diversion. The proof essentially rests on two observations: (1) For $t=0$, $XM(0,\gamma_F)=0$; and (2) under Condition 1, $XM(0,\gamma_F)$ is decreasing and concave. The first observation is obvious, so let t be positive. For relatively efficient relocations, x is near 0. The welfare effect $r(0,t)=\Delta\Psi\left(\widetilde{\omega}_\beta,t\right)$ may be positive or negative, but $r(x,t)$ falls as x increases. Now, if the flow rate of new matches with productivity near $\widetilde{\omega}_\beta$ is the same as the flow rate of new matches with suppliers with productivity near $\widetilde{\omega}^P_F$, then the negative effects due to the latter group of rematches will dominate and make $XM(0,\gamma_F)<0$. Under Condition 1, the flow rate is non-decreasing in the tariff. Hence, the negative effects receive higher weight than the (possibly) positive effects. It follows that $XM(0,\gamma_F)$ is decreasing and concave in t.

Now we return to the consideration of the full welfare effect, and show that the external tariff that maximizes welfare for a large North-South PTA partner ($\gamma_F=1$) is inefficiently high for a smaller North-South partner ($\gamma_F<1$). The tariff preference has a better effect when ω is lower. Thus, to maximize the aggregate within-chain effect, it is optimal to have an external tariff that promotes a high enough relationship-strengthening effect for the best suppliers even when that comes at the cost of lowering the welfare created by the marginal incumbent supplier. Hence, $\Delta\Psi(\widetilde{\omega}_\beta,t)=r(0,t)$ is decreasing in t at $t=\hat{t}_W^{\gamma_F=1}$ and welfare from all relocations falls with t. This also implies that the range of tariffs such that the PTA enhances welfare is smaller when $\gamma_F<1$. We have the following.

Proposition 11 *For any $\gamma_F<1$, the welfare effect of a North-South PTA, $\Delta W(0,\gamma_F)$, is maximized for some $t\in(0,\hat{t}_W^{\gamma_F=1})$. Moreover, there exists a $\underline{t}>0$ such that if $t<\underline{t}$, then a North-South PTA enhances aggregate welfare. Also, there exists a $\bar{t}\in[\underline{t},2\hat{t}_W^{\gamma_F=1})$ such that if $t>\bar{t}$, then a North-South PTA lowers aggregate welfare. Under Condition 1, $\underline{t}=\bar{t}$ is unique.*

Proof. Note that if $t=0$, then $\Delta W(0,\gamma_F)=0$. Define $\underline{t}$ to be the lowest non-negative value of t such that $\Delta W(0,\gamma_F)=0$. Differentiating, we have that $\frac{d\Delta W(0,\gamma_F)}{dt}=\frac{dIM(0,\gamma_F)}{dt}+\frac{dXM(0,\gamma_F)}{dt}$. In the limit, $\lim_{t\to 0}\frac{dIM(0,\gamma_F)}{dt} > \lim_{t\to 0}\frac{dXM(0,\gamma_F)}{dt} = 0$. Hence, $\lim_{t\to 0}\frac{d\Delta W(0,\gamma_F)}{dt} > 0$ and $\underline{t} > 0$. Hence, $t = 0$ does not maximize $\Delta W(0,\gamma_F)$.

From Proposition 7, $IM(0,\gamma_F)$ is maximized at $t = \hat{t}_W^{\gamma_F=1} > \widehat{t}(\widetilde{\omega}_\beta)$. We now show that $XM(0,\gamma_F) = \int_0^t r(x,t)\phi(x;\gamma_F,G)dx$ is decreasing in t at $t = \widehat{t}(\widetilde{\omega}_\beta)$. Note first that $r(0,t) = \Delta\Psi(\widetilde{\omega}_\beta, t)$ is maximized at $t = t(\widetilde{\omega}_\beta)$. By Lemma 4, it follows that $r(0,t)$ decreases in t for any $t > t(\widetilde{\omega}_\beta)$. We can also show that $\frac{dr(x,t)}{dt}$ is decreasing in x:

$$\frac{d^2 r(x,t)}{dxdt} = \frac{-4(1-\gamma_F)\alpha(1-\alpha)b^2}{\left(2c-\alpha b^2\right)^2} < 0.$$

This implies that if $t > t(\widetilde{\omega}_\beta)$, then

$$\int_0^t \frac{dr(x,t)}{dt}\phi(x;\gamma,G)dx < 0.$$

Because $\phi(t;\gamma,F)r(t,t) < 0$ (Lemma 5), we have that for any $t \geq \hat{t}_W^{\gamma_F=1}$,

$$\frac{dXM(0,\gamma_F)}{dt} = \phi(t;\gamma,G)r(t,t) + \int_0^t \frac{dr(x,t)}{dt}\phi(x;\gamma,G)dx < 0.$$

This shows that $XM(0,\gamma_F)$ is decreasing in t for any $t \geq \hat{t}_W^{\gamma_F=1}$ and therefore $t \geq \hat{t}_W^{\gamma_F=1}$ does not maximize $\Delta W(0,\gamma_F)$. So restrict attention to $t \in [0, \hat{t}_W^{\gamma_F=1}]$, a compact set. Because $\Delta W(0,\gamma_F)$ is continuous, it is maximized somewhere for some t in that set. Having ruled out the boundaries, we conclude it is maximized for some $t \in (0, \hat{t}_W^{\gamma_F=1})$.

From Proposition 9, $XM(0,\gamma_F) < 0$ for any $t > 2\widehat{t}(\widetilde{\omega}_\beta)$ and $\hat{t}_W^{\gamma_F=1} > \widehat{t}(\widetilde{\omega}_\beta)$. Hence, if $t \geq 2\hat{t}_W^{\gamma_F=1}$, then $\Delta W(0,\gamma_F) = IM(0,\gamma_F) + XM(0,\gamma_F) < 0$. By continuity of $\Delta W(0,\gamma_F)$, it follows that $\Delta W(0,\gamma_F) < 0$ for some $t < 2\hat{t}_W^{\gamma_F=1}$ as well. Define $\bar{t}$ to be the highest t such that $\Delta W(\gamma) = 0$. Thus, we have shown that $\bar{t} \in [\underline{t}, 2\hat{t}_W^{\gamma_F=1})$.

Finally, Condition 1 implies that $\Delta W(0,\gamma_F)$ is strictly concave in t. Hence, $\Delta W(0,\gamma_F) = 0$ for just one value of $t = \underline{t} = \bar{t}$. ■

Under Condition 1, the extensive-margin welfare effect is strictly concave; hence, $\Delta W(0,\gamma_F)$ is strictly concave as well and there is a uniquely optimal t and a single crossing of zero. If Condition 1 does not hold, there may be (with respect to t) multiple local maxima and multiple crossings of zero.