Preferential Trade Agreements and Global Sourcing

Emanuel Ornelas, John L. Turner, Grant Bickwit

Preferential Trade Agreements and Global Sourcing

Abstract

We study how a preferential trade agreement (PTA) affects international sourcing decisions, aggregate productivity and welfare under incomplete contracting and endogenous matching. Contract incompleteness implies underinvestment. That inefficiency is mitigated by a PTA, because the agreement allows the parties in a vertical chain to internalize a larger return from the investment. This raises aggregate productivity. On the other hand, the agreement yields sourcing diversion. More efficient suppliers tilt the tradeoff toward the (potentially) beneficial relationship-strengthening effect; a high external tariff tips it toward harmful sourcing diversion. A PTA also affects the structure of vertical chains in the economy. As tariff preferences attract too many matches to the bloc, the average productivity of the industry tends to fall. When the agreement incorporates "deep integration" provisions, it boosts trade flows, but not necessarily welfare. Rather, "deep integration" improves upon "shallow integration" if and only if the original investment inefficiencies are serious enough. On the whole, we offer a new framework to study the benefits and costs from preferential liberalization in the context of global sourcing.

Emanuel Ornelas
Sao Paulo School of Economics-FGV
Sao Paulo / Brazil
eaornelas@gmail.com

John L. Turner
Department of Economics
University of Georgia
Athens / GA / USA
jlturner@uga.edu

Grant Bickwit
University of Virginia
Darden School of Business
Charlottesville / Virginia / USA
BickwitG19@darden.virginia.edu

1 Introduction

The past few decades have seen 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 last 40 years 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 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 trade in customized inputs. This is what we aim to provide in this paper. We consider a market with endogenous formation of two-firm vertical chains and non-contractible investments that are specific to relationships within each chain. We show that PTAs can be welfare-improving even if conventional "trade creation" forces are absent, because tariff preferences serve as an (imperfect) substitute for complete contracts and stimulate value creation within chains. This is especially true for high-productivity industries. But tariff preferences also yield production of too many specialized inputs, and induce the destruction of high-productivity chains outside the PTA in exchange for low-productivity chains inside the bloc. The implications for "deep integration" are also entirely novel: deep provisions are helpful only when original inefficiencies are sufficiently severe, but not otherwise.

Our model therefore contrasts with standard regionalism theories in its motivation, its mechanisms and its results. Since Viner (1950), analyses of preferential liberalization have typically pointed to two opposing effects of preferential tariffs, trade creation and trade diversion. Trade creation occurs when firms from foreign partner countries produce more due to the PTA, at the expense of inefficient domestic firms. This increases overall welfare. Trade diversion occurs when member-country firms produce more due to the PTA, but at the expense of efficient nonmember firms. This lowers overall welfare. Those effects are based upon classical trade models, which rely

---

[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).

on market clearing for price formation. That is also the approach taken in modern quantitative analyses of the welfare implications of PTAs, such as Caliendo and Parro's (2015). While they take trade in intermediate products explicitly into account, their model is based on comparative advantage forces, with anonymous markets and well-defined world prices for all goods.

In reality, modern trade in intermediates often involves customized components that commit a buyer and a seller to each other. First, they need to find each other. Once matched, they become locked in to each other and may underinvest in component-specific technology due to 'hold-up problems' when contracts are incomplete (e.g., as in Grossman and Hart, 1986). For example, a buyer of customized components can hold up the seller and force a new bargain where he captures some of the surplus created by sunk investments made by the seller. As the seller anticipates that outcome, she underinvests.

We introduce a property-rights model coupled with a Walrasian matching process to capture those effects in as simple way as possible. Suppliers in different countries and with different levels of productivity match with buyers to form vertical chains. Each supplier customizes her inputs to the buyer within their chain, and they bargain over terms of trade. Each buyer may source customized inputs from within his chain and/or generic inputs from a competitive market.[2] The PTA affects matching, customization investments and the composition of sourced inputs. Importantly, we design the model to shut down all Vinerian trade creation channels. We put aside classic trade creation not because we deem it unimportant,[3] but to shed light on potentially important forces that have so far been ignored in the academic literature and in policy circles alike.

In our 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 former group, which we call *incumbent suppliers*, the responses to preferential access generate a positive welfare effect if and only if the external tariff is sufficiently low, and the welfare effect is higher whenever the distribution of supplier productivity is better, in the sense of stochastic dominance. For the latter group, which we call *new suppliers*, the welfare effect is more nuanced because the distribution of supplier productivity itself changes. Since new suppliers are less productive than those they replace, and since the firms do not internalize the full

---

[2]We use the term *Y-chain* to describe the entire supply chain. See Figure 1 on p. 10.

[3]After all, as Freund and Ornelas (2010) conclude from the existing literature, trade creation seems to be more prevalent than trade diversion in actual PTAs.

welfare consequences of rematching, the range of tariffs such that the total welfare effect of the PTA is positive is smaller when there are new suppliers. Still, there are tariff levels and productivity distributions such that the emergence of new suppliers enhances welfare over and above the effect generated by incumbent suppliers.

To understand the mechanisms, it is instructive to consider first the impact for incumbent suppliers. Under a PTA, they receive a higher surplus on every unit traded. This propels more trade in customized 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 is necessarily positive when the external tariff is low, but a sufficiently high external tariff induces an excess of investment. On the other hand, there is the usual negative effect from tariff discrimination—here, trade diversion in the sourcing of components, from generics to expensive customized inputs—which increases monotonically in the tariff. 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 initially trade high volumes—i.e., when they have high productivity.

For incumbent suppliers, the welfare effect of the PTA is determined entirely by those two forces. When external tariffs are very low, PTAs raise welfare for sure. In contrast, if external tariffs are sufficiently high, PTAs are likely to harm welfare. Thus, as in the classical case, with very high preferential tariffs, trade diversion dominates. Yet recall that here the comparison is not with classic trade creation, but with the relationship-strengthening effect. When tariff preferences are too high, they yield "too much" investment, more than offsetting the benefit of alleviating the original hold-up problem. The welfare effect is also higher when incumbent suppliers are more productive. Hence, we introduce a new element into Viner's classic tradeoff by showing that tariff preferences are more likely to enhance welfare when applied to more efficient industries, which trade large volumes of specialized inputs even without the PTA.[4]

---

[4] This result is reminiscent of the "natural trading partners" hypothesis, which posits that agreements formed between countries that trade heavily with each other are more likely to enhance welfare. 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 (e.g., Bhagwati and Panagariya, 1996). Our result provides a possible rationale for it.

Consider next new suppliers. A domestic buyer matched with a supplier in a non-PTA country can earn higher profit by matching with a supplier with the same productivity in a PTA country. When a PTA is formed, some buyers then break chains with existing suppliers outside the PTA and form chains with PTA insiders. Once rematched, they benefit both from the tariff preference and from the improved investment incentives of the new suppliers. Two intuitive economic forces push welfare in a negative direction. First, suppliers lost outside the PTA are (pre-investment) more productive than those gained inside the PTA. Second, the marginal chains formed are unambiguously bad for welfare, in spite of the new investments. The reason is that matches are based on private profits and fail to internalize lost tariff revenue.

Still, the new supplier effect on welfare *can* be positive. Two conditions are needed for that. First, all incumbent suppliers must yield welfare gains under the PTA. Second, the mass of new suppliers must be relatively similar to the least-productive incumbent supplier, so that the fundamental productivity of the industry does not deteriorate much with the agreement.

Observe that the mechanisms behind our results affect not only allocative inefficiency (as e.g. in Antràs and Staiger, 2012a). Here, PTAs also yield changes in the *production process* and in the *formation of vertical chains*, both of which affect the aggregate productivity of the economy. All of that happens simply because of the tariff preference. The upshot is that the welfare implications of PTAs under global sourcing are much more subtle and intricate than standard models suggest.

This becomes even more evident when we model deep integration features of PTAs, like stronger bilateral recognition of intellectual property rights. We show that they have a positive effect on trade flows, in line with the empirical literature (e.g., Mattoo, Mulabdic and Ruta, 2017), but not necessarily on welfare. Whether deep integration is helpful or not will depend on pre-agreement inefficiencies in investment. It follows that some countries may actually be better off if they kept their agreements "shallow."

Thus, 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 specialized inputs, tariff preferences are not just policy instruments that directly affect prices; they also affect the efficiency of the production process, through changes in the incentives to invest and to form

4

vertical chains.

In that sense, our paper adds to the literature that seeks to link 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 model 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, our emphasis is precisely on the discriminatory aspect of tariff changes. Furthermore, we are interested in how they affect investment and matching patterns related to international sourcing decisions, not a special concern in the analyses of Bustos (2011) and Lileeva and Trefler (2010).

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 example, as in much of the international sourcing literature, the role of market thickness in shaping outsourcing decisions feature prominently in Grossman and Helpman (2005), whereas we concentrate on the themes described above.

In terms of structure, we build on Ornelas and Turner (2008, 2012), but pursue very different directions. Our previous papers study neither preferential liberalization, our focus here, nor deep integration, and do not consider heterogeneity in productivity and endogenous matching, both essential ingredients of the current analysis.

The paper is also closely related to Antràs and Staiger (2012a, b). Although their goal is to study the optimal design of (nondiscriminatory) trade agreements, not an issue we address, their more general point is that the efficiency properties of international trade agreements are vastly

---

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

different when buyers/consumer 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 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), allowing us to generate very different results. In particular, unlike in their setting, we underscore how tariff preferences shape the structure of the production process through their effects on investment and matching decisions.[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 and study the equilibrium without a trade agreement in section 3. In section 4 we analyze the equilibrium with a PTA and describe its impact on firms' choices. We then assess the welfare impact of the PTA in section 5. In section 6 we discuss the robustness of our results to alternative specifications, and we extend the analysis to trade agreements with "deep integration" features in section 7. Finally, we discuss some testable implications of our model in section 8 and conclude in section 9.

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. There is also a numéraire good y that enters consumers' utility function linearly. Thus, if consumers 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,

---

[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 (2015) study as well how trade policy affects international patterns of procurement, but their proposed mechanism—how changes in trade policy uncertainty affects the mode of sourcing relationships— is very different from ours. From a different angle, Antràs and de Gortari (2017) 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.

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 process and sell 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. Entirely analogous analyses could be carried out for other differentiated sectors.

There is another country, *Foreign*, as well as the rest of the world (ROW). When sourcing, each buyer may purchase generic inputs g available in the world market and/or customized inputs q from a specialized supplier (S). Specialized suppliers are located in either *Foreign* or *ROW*. 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. For expositional simplicity, we assume that neither *Home* nor *Foreign* produces generic inputs. This is not without loss of generality, but helps us convey our main ideas in the simplest possible way. In section 6 we discuss how alternative configurations of the generic industry would affect our results.

Home's buyers face a per-unit tariff t on all imported intermediate goods, so a generic input costs $p_w + t$ for them. Generally, a buyer values generic and customized inputs differently. However, we can define units so that one unit of generic input and one of customized input have the same revenue-generating value for a buyer.[7] Under this normalization, all that matters for B's revenue is the total number of intermediate inputs he purchases, $Q = g + q$, not the composition of Q.[8]

Now, to acquire customized 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*. Suppliers are split between *Foreign* and *ROW* proportionally to γ and $1 - \gamma$, respectively. We assume that $\beta < \gamma$. This implies that buyers would remain scarce relative

---

[7] For example, we could 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 g (adjusted for quality) is not essential, but it is critical that they are substitutes to some degree.

to suppliers even if they matched only in *Foreign*. Each supplier is identified by ω, a heterogeneity parameter that indexes (the inverse of) her productivity. The distribution of suppliers in each country follows distribution $F(\omega)$, with an associated density $f(\omega)$, where ω lies on $[0, p_w]$.[9] To focus on fundamental forces, we consider the simplest possible matching framework, namely a Walrasian 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.

Upon forming a vertical chain, B and S specialize their technologies toward each other. This specialization costs nothing, but implies that at any point in time 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.

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.[10] 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 specialized inputs from S. Finally, B transforms all inputs into the final good and payoffs are realized.

In order to generate clear-cut analytical solutions, we adopt some specific functional forms. Conditional on investment i, we specify the supplier's cost function as

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

where q denotes her customized input production. Parameter ω shifts the firm's marginal cost; the

---

[9]As it will become clear shortly, in the absence of trade agreements 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 j with $\omega_j \geq p_w$ 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 $F(\omega)$ is the truncated distribution of suppliers when $\omega \leq p_w$.

[10]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.

lower is ω, the more efficient the firm is. In turn, c determines the slope of the supplier's marginal cost, while b denotes the effectiveness of investment in reducing her production costs. In turn, the cost of the investment is

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

Investment is bounded by $i \in [0, i^{max}]$. We assume that $2c > b^2$.[11]

Concrete functional forms are useful to analyze PTAs, where changes in tariffs are not marginal but discrete, from their initial levels to zero, and where we want to condition results on the extent of the margin of preference. The linear-quadratic specification that we adopt displays properties that are standard and provide a good representation of the key elements of our environment: investment and original 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); 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.

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 require that C_{qqq} and I''' should be sufficiently small in absolute value, but the analysis becomes particularly clean if one sets them to zero, as we do here.

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, we assume $C_q(Q^*, i^{max}, 0) > p_w$, so that even under the maximum investment (and under free trade), the marginal cost for the most productive firm ($\omega = 0$) is still sufficiently high that B prefers to purchase some generic inputs. In addition to being realistic,[12] the main role of the dual sourcing

---
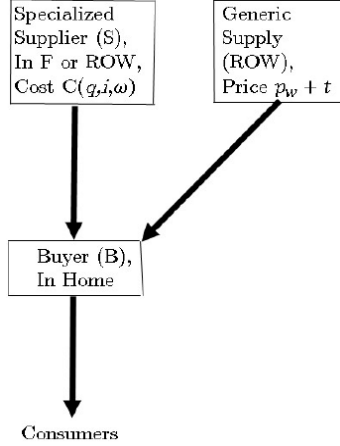
Fig. 1: Y-chain

specification is pedagogical, as will become clear in the analysis. 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.

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. The timing of events is summarized as follows:

- Each B matches with a supplier S in either *Foreign* or *ROW* to form a vertical chain, adapting their technologies toward each other within the chain;

- S makes an irreversible relationship-specific investment;

- 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 only generic inputs;

- B purchases g;

- Final production occurs and final goods are sold.

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

10

3   No Trade Agreement

When there is no trade agreement, all inputs imported into *Home* are subject to the tariff regardless of their origin.

3.1   Single Partnership

After S chooses her investment, B and S determine the price of the specialized intermediate inputs, p_N^s, by Generalized Nash Bargaining over the surplus due to trading q_N customized inputs instead of only generic ones. Specifically, let the supplier have bargaining power $\alpha \in (0,1)$. Under Generalized Nash Bargaining, the two firms choose p_N^s to maximize

$$(U_B^T - U_B^0)^{(1-\alpha)}(U_S^T - U_S^0)^\alpha,$$

where U_k^J is the verifiable profit that firm k (either B or S) would receive under scenario J. The two possible scenarios are either bargaining and trading (T) or not reaching an agreement and thus not trading (0). Those values are laid out as follows: $U_B^T = V(Q^*) - (p_w + t)g_N - (p_N^s + t)q_N$; $U_B^0 = V(Q^*) - (p_w + t)Q^*$; $U_S^T = p_N^s q_N - C(q_N, i, \omega)$; $U_S^0 = 0$.

Defining $\Omega \equiv (U_B^T - U_B^0) + (U_S^T - U_S^0)$ as the bargaining surplus, the outcome of bargaining has the two firms splitting the proceeds, with S receiving $\alpha\Omega$ and B receiving $(1-\alpha)\Omega$, in addition to their reservation payoff, U_k^0. In the absence of a trade agreement,

$$\Omega_N = p_w q_N - C(q_N, i_N, \omega). \tag{2}$$

Conditional on investment i and on the tariff, a B-S vertical chain trades the ex-post privately optimal number of specialized inputs, q_N, and B purchases the ex-post privately efficient level of generic inputs, g_N. Together, they compose the total number of inputs purchased within the Y-chain by B: $Q^* = q_N + g_N$. Since without a PTA both customized and generic inputs incur the tariff, privately optimal sourcing equalizes the marginal cost of the two alternative inputs,

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

pinning down q_N (and hence g_N) for given i. Under our functional form specification, this condition becomes

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

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

$$\max_{i_N} \alpha \Omega_N - I(i_N).$$

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

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

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

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

Hence, the equilibrium investment and output are proportional. More productive (lower-ω) firms produce more for a given investment, and they also invest more, reinforcing their original advantages. When the supplier's bargaining power (α) is very small, the investment is very low, and drops to zero as $\alpha \to 0$, when S does not appropriate any of the benefits of her investment. As α rises, 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 also that neither investment nor production is affected by the tariff, which in this setting distorts the total volume of inputs, Q^*, but does not interfere with the sourcing of q.

It is useful to compare S's investment choice with the efficient level of investment, given the tariff. Under privately efficient sourcing, worldwide social welfare due to this bilateral relationship can be defined as

$$\Psi_N = V(Q^*) - p_w Q^* + p_w q_N - C(q_N, i, \omega) - I(i). \tag{7}$$

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 functional form specification, this yields

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

Observe that, as b approaches $\sqrt{2c}$, the level of the efficient investment blows up.[13] Comparing i_N^* with i^e, it is immediate that $i_N^* < i^e$ (since $\alpha < 1$). Moreover, it is easy to see that the extent of the hold-up problem, which we can define as $HUP_N \equiv i^e - i_N^*$, depends critically on the productivity of the supplier:

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

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

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

which is clearly decreasing in ω. ∎

Intuitively, this happens because actual investment increases with S's share α of the bargaining surplus, whereas the efficient level of investment increases with the whole bargaining surplus. The extent of the inefficiency is therefore proportional to $(1 - \alpha)\,\Omega_N$, but Ω_N is itself increasing in productivity. Hence, it is precisely the vertical chains with the best suppliers—who produce more and generate higher surplus for any level of investment—that are more negatively affected by contract incompleteness.

---

[13] In this case, $i^{\max}$ would obtain as a corner solution.

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

$$U_S^N(\omega) = \frac{\alpha (p_w - \omega)^2}{2c - \alpha b^2}, \tag{10}$$

$$U_B^N(\omega) = \frac{2c(1 - \alpha)(p_w - \omega)^2}{(2c - \alpha b^2)^2}. \tag{11}$$

Both are clearly decreasing in ω, so low-ω suppliers earn higher profits than high-ω suppliers, and a buyer's profit is higher in a vertical chain with a low-ω supplier.

3.2 Structure of Matches

Initially, suppliers and buyers are not specialized to each other. Each B matches with a supplier S in either *Foreign* or *ROW* to form a vertical chain. We consider a Walrasian matching environment where each supplier that 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 their current matches and forming a new match with a new mutually-agreeable fee. Hence, we can use the intuitive logic of stability to help describe equilibrium.

Feasibility requires that the measure of suppliers matched cannot 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. Hence, denoting the hypothetical values for the cutoff levels of productivity in *Foreign* and *ROW* by $\widehat{\omega}_F$ and $\widehat{\omega}_{ROW}$, respectively, in a feasible equilibrium we must have the following market-clearing condition:

$$\gamma \int_0^{\widehat{\omega}_F} dF(\omega) + (1 - \gamma) \int_0^{\widehat{\omega}_{ROW}} dF(\omega) = \beta. \tag{12}$$

Additionally, the marginal matches in *Foreign* and *ROW* must yield the same joint payoff to the members of the partnership. As the distribution of suppliers is the same in the two markets, and the joint future payoff of a B-S chain for a given ω is also equal in both markets in the absence

14

of trade agreements, in equilibrium the marginal matches in each market must involve suppliers with the same productivity:

$$\widehat{\omega}_F = \widehat{\omega}_{ROW}. \tag{13}$$

Using those two conditions, we then have that equilibrium in the market for matches without a PTA implies $\widehat{\omega}_F = \widehat{\omega}_{ROW} = \widetilde{\omega}_N$, where $\widetilde{\omega}_N$ is determined by

$$F(\widetilde{\omega}_N) = \beta. \tag{14}$$

Observe that a larger *Home* (i.e., a higher β) implies a higher cutoff $\widetilde{\omega}_N$, with buyers matching further down in the productivity distribution. The relative size parameter γ does not affect the distribution of productivity among suppliers that match.

Because all buyers are identical, each supplier is indifferent about the buyer to whom it is matched and cares only about the size of the fee paid. Buyers care about both the size of the fee and the supplier's productivity, which affects the buyer's ultimate profit. Equilibrium is achieved when each buyer earns the same profit, so the fee must differ across matches. To see why, suppose that there is just one fee. Then a buyer matched to a relatively low-productivity supplier would earn a relatively low profit. He would prefer to match for a lower fee with a higher-productivity supplier, and the higher-productivity supplier would also prefer this.

Hence, the fee paid to a buyer must depend upon the productivity of its matched supplier. Specifically, the equilibrium matching fee schedule is the same for matches with suppliers in *Foreign* and *ROW,* and satisfies

$$M_N(\omega) = U_S^N(\widetilde{\omega}_N) - \left[U_B^N(\omega) - U_B^N(\widetilde{\omega}_N) \right].$$

Note that all buyers earn $U_B^N(\omega) + M_N(\omega) = U_B^N(\widetilde{\omega}_N) + U_S^N(\widetilde{\omega}_N) > 0$, so their payoffs are invariant to ω. This happens because, as a higher productivity of the matched supplier increases $U_B^N(\omega)$, the buyer's fee decreases by exactly the same amount. In contrast, the cutoff supplier earns a payoff of exactly 0 but higher-productivity suppliers earn more, as they absorb the whole extra aggregate surplus brought about by the higher productivity through a lower fee to the buyers.

4   A Preferential Trade Agreement

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*.

For vertical chains with suppliers in *ROW* before and after the PTA, the previous analysis applies in its entirety; the changes are restricted to vertical chains with suppliers in *Foreign* and to those where the buyer decides to change the location of his match. Since generic inputs remain imported from *ROW*, they still cost $p_w + t$ for *Home*'s buyers.

4.1   Single Partnership

Consider a vertical chain with a supplier located in *Foreign*. 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 changes to reflect the new relative prices. This is summarized by

$$C_q(q_P, i_P, \omega) = p_w + t, \tag{15}$$

which under our functional form specification is equivalent to

$$q_P = \frac{p_w + t - \omega + bi}{c}. \tag{16}$$

Only one of the potential U_k^J payoff terms, U_B^T, structurally changes, becoming

$$U_B^T = V(Q^*) - (p_w + t)g_P - p_P^s q_P.$$

The bargaining surplus under a trade agreement, Ω_P, is defined in the same manner as before, but now reflects the change in buyer profit with trade due to tariff savings when B sources from S:

$$\Omega_P = (p_w + t)q_P - C(q_P, i_P, \omega).$$

Due to Generalized Nash Bargaining, B and S retain the same shares of Ω_P as they do without a trade agreement. Accordingly, the investment decision is conceptually unchanged, being the

solution of

$$\max_{i} \alpha\Omega_P - I(i_P).$$

The equilibrium level of investment under the PTA can then be expressed as

$$i_P^* = \left(\frac{\alpha b}{2c - \alpha b^2}\right)(p_w + t - \omega). \tag{17}$$

Clearly, the preferential trade agreement induces an increase in relationship-specific investments.

We define the change in investment due to the PTA as $\Delta i \equiv i_P^* - i_N^*$. Our quadratic specification yields the useful property that it is proportional to the tariff:[14]

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

The change in investment vanishes when $\alpha \to 0$ and is strictly increasing (at an increasing rate) in α. It also increases with the responsiveness of marginal cost to investment (b) and decreases with the slope of the marginal cost curve (c).

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

$$q_P^* = \left(\frac{2}{\alpha b}\right)i_P^*, \tag{18}$$

and therefore the effect of the PTA on the number of customized inputs, $\Delta q \equiv q_P^* - q_N^*$, also is 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 advantage from not facing the tariff. This effect takes place even if there were no additional investment. In particular, observe that if the investment did not lower production cost ($b = 0$), the supplier would never invest and yet sales of customized inputs would still increase, by $\Delta q(b = 0) = \frac{t}{c} > 0$.

_____

[14]The multiplicative constant in Δi is analogous to what we termed the "investment effect" of a tariff in our previous work in the context of nondiscriminatory liberalization (Ornelas and Turner 2008; 2012).
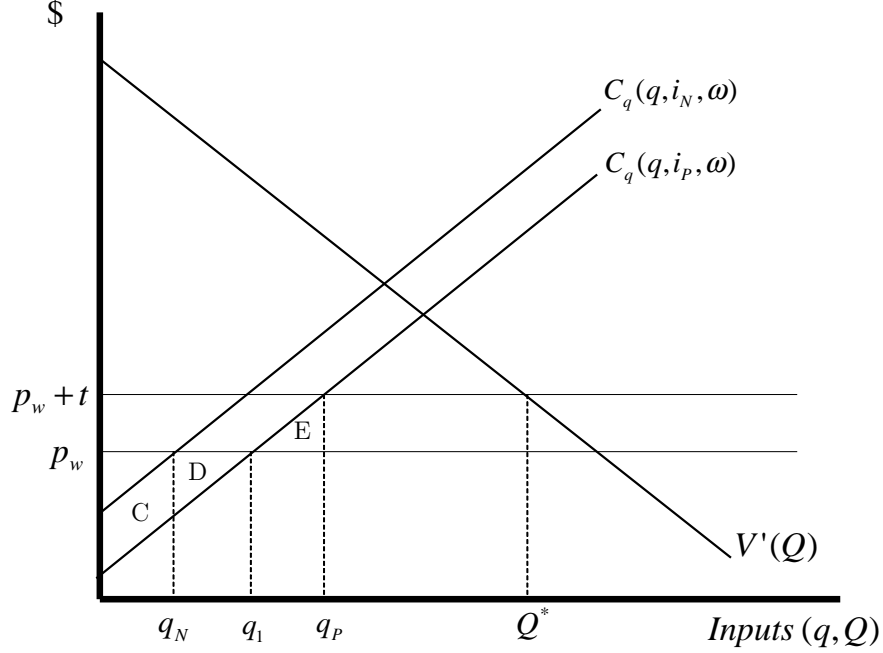
Fig. 2: The Effects of a PTA on Sourcing and Production

The sales of specialized inputs increase also because of lower production costs. Under the PTA, S's investment enhances the bargaining surplus by more than it does without a trade agreement. Since S keeps some of those gains, she has an incentive to increase her investment. When investment is higher, S's entire marginal cost curve is lower. There are then more units that, from an efficiency standpoint, *should* be produced by S. Such level, q_1^*, satisfies $C_q(q_1^*, i_P, \omega) = p_w$. Developing this expression under our functional form specification and using (4), we obtain

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

It is easy to see that

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

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

Figure 2 highlights the effects of the PTA within a single Y-chain. Units $q \in (0, q_N)$ 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 $q \in (q_N, q_1)$ 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_P. They reflect classic trade diversion, as the extra customized 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, $HUP_N = i^e - i_N^* > 0$ for sure. Such an unambiguous ordering does not exist under the PTA. Defining the excess of investment under a PTA as $EXC_P \equiv i_P^* - i^e$,[16] one finds that

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

It follows that $i_P^* > i^e$ when α is sufficiently close to one (in which case the original hold-up problem is relatively unimportant, so the investment boost due to the PTA is mostly distortionary) and/or when t is sufficiently high (in which case the PTA is too effective in encouraging investment).

Overall, this analysis highlights a "within Y-chain" tradeoff between conventional trade/sourcing diversion and an effect that so far has been entirely neglected in the regionalism literature. Due to the PTA, the chain creates additional surplus for all units of customized inputs that would be produced without the agreement, plus some surplus for additional units traded—areas **C** and **D** in Figure 2. This increases welfare, possibly more than offsetting the losses due to excessive production (area **E**) and additional investment.

It is important to stress at this point that, while our model displays an effect akin to Vinerian trade diversion, *Vinerian trade creation is shut down.* Classic trade creation would be observed if the PTA led to more total units traded, but Q^* is kept fixed by design (for given t). Thus, if one considered only traditional forces, one would deem the model designed to highlight the negative

---

[15]Observe that a change in parameter ω provokes a parallel shift of the marginal cost curve. It is easy to see that such 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 XXX), 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_N^*.

[16]In the Appendix we show that the efficient level of investment is the same under no agreement and under a PTA.

welfare consequences of PTAs. Instead, it is designed to shed light on novel channels through which PTAs affect economic efficiency.

With a PTA, we can solve for closed-form expressions for equilibrium profits conditional on ω:

$$
\begin{aligned}
U_S^P(\omega, t) &= \frac{\alpha (p_w + t - \omega)^2}{2c - \alpha b^2}, && (19) \\
U_B^P(\omega, t) &= \frac{2c(1 - \alpha)(p_w + t - \omega)^2}{(2c - \alpha b^2)^2}. && (20)
\end{aligned}
$$

Again, both are clearly decreasing in ω.

Consider next a vertical chain with a supplier in ROW. As stated earlier, the "no PTA" analysis applies in its entirety. The profits of the supplier and buyer are the same as in (10)-(11). Note that these payoffs are the same as in (19)-(20) with $t = 0$, i.e., $U_S^N(\omega) = U_S^P(\omega, 0)$ and $U_B^N(\omega) = U_B^P(\omega, 0)$. We will generally use the $U_S^N(\omega)$ and $U_B^N(\omega)$ expressions when referring to profits from vertical chains with suppliers in ROW. More generally, whenever we drop the t argument from a function, that means that there is no discriminatory protection $(t = 0)$ and equilibrium outcomes follow the "no PTA" case.

4.2 Structure of Matches

Analogously to section 3.2, we first describe the characteristics of the competitive matching equilibrium and then discuss how the equilibrium is achieved. The market-clearing condition (12) is unchanged with the PTA. And once again it suffices to identify a condition requiring that the marginal matches in *Foreign* and *ROW* yield the same joint payoff to the members of the vertical chain. However, when *Home* forms a PTA with *Foreign*, a supplier with productivity ω generates a higher aggregate payoff if she is located in *Foreign*. Simple inspection of (10), (11), (19) and (20) makes clear that[17]

$$
\widehat{\omega}_F = \widehat{\omega}_{ROW} + t. \tag{21}
$$

Using conditions (12) and (21), we then have that equilibrium in the market for matches under

---

[17]If the external tariff were sufficiently high, we would have $\Omega_P^*(\omega + t) > \Omega_N^*(\omega)$ for all $\omega \geq 0$. In that case, all buyers would match with suppliers in *Foreign* and $\widetilde{\omega}_{ROW}$ would be undefined. Qualitatively, the analysis would be very similar, but to avoid a taxonomy we concentrate on the case where there are matches in both locations.

a PTA implies

$$\gamma F(\widetilde{\omega}_{ROW} + t) + (1 - \gamma)F(\widetilde{\omega}_{ROW}) = \beta. \qquad (22)$$

This determines $\widetilde{\omega}_{ROW}$. Using (21), we obtain $\widetilde{\omega}_F$.

It is straightforward to see that $\widetilde{\omega}_N \in (\widetilde{\omega}_{ROW}, \widetilde{\omega}_F)$. Hence, when *Home* forms a PTA with *Foreign*, some buyers that would have matched with suppliers in *ROW* that are more productive than $\widetilde{\omega}_N$ end up matched with suppliers in *Foreign* that are less productive than $\widetilde{\omega}_N$. This difference is maximal when we consider the hypothetical 'last' buyer to switch suppliers, who leaves a supplier with productivity $\widetilde{\omega}_{ROW}$ in *ROW* for a supplier with productivity $\widetilde{\omega}_F$ in *Foreign*. Both matches yield the same aggregate payoff for the vertical chains, as the difference in productivity between them is exactly offset by the (direct and indirect) benefits from the tariff preference.

The equilibrium matching fee schedules for matches with suppliers in *Foreign* and *ROW* satisfy

$$
\begin{aligned}
M_{P,ROW}(\omega) &= U_S^N(\widetilde{\omega}_{ROW}) - \left[U_B^N(\omega) - U_B^N(\widetilde{\omega}_{ROW})\right], \\
M_{P,F}(\omega) &= U_S^P(\widetilde{\omega}_F, t) - \left[U_B^P(\omega, t) - U_B^P(\widetilde{\omega}_F, t)\right].
\end{aligned}
$$

As with equilibrium under no PTA, all buyers earn the same payoff of $U_S^N(\widetilde{\omega}_{ROW}) + U_B^N(\widetilde{\omega}_{ROW})$. This is higher than the payoff of $U_S^N(\widetilde{\omega}_N) + U_B^N(\widetilde{\omega}_N)$ that buyers earn under no PTA. Once again, the cutoff suppliers earn a payoff of zero, while higher-productivity suppliers earn positive profits. The most profitable supplier is the $\omega = 0$ supplier in *Foreign*.

Figure 3 illustrates the matching equilibrium. It shows equations (12), (13) and (21) for hypothetical values of the cutoff levels of productivity in *Foreign* and *ROW*, $\widehat{\omega}_F$ and $\widehat{\omega}_{ROW}$. The equilibrium cutoff $\widetilde{\omega}_N$ satisfies (12) and (13) for the no-PTA case, while $\widetilde{\omega}_F$ and $\widetilde{\omega}_{ROW}$ satisfy (12) and (21) for the PTA case. The downward-sloping function is implied by (12). As $\widehat{\omega}_{ROW}$ increases, there are more vertical chains formed with suppliers in *ROW*. Hence, the number of vertical chains formed with suppliers in *Foreign* must fall. When $\widehat{\omega}_{ROW} = \widehat{\omega}_F$, it follows that $F(\widehat{\omega}_{ROW}) = \beta$, so this yields $\widetilde{\omega}_N$.

Comparative statics follow directly from the figure. A higher external tariff t shifts equation (21) upwards. This increases the PTA cutoff in *Foreign*, $\widetilde{\omega}_F$, and decreases the PTA cutoff in *ROW*, $\widetilde{\omega}_{ROW}$. Intuitively, a higher tariff drives a bigger wedge between the productivities of the
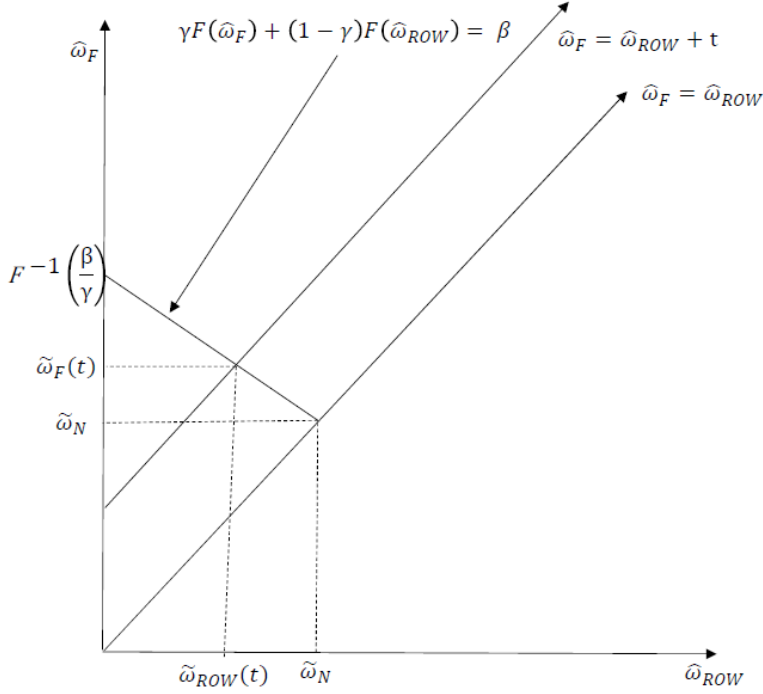
Fig. 3: Matching Equilibrium with and without the PTA

suppliers in the marginal re-match. The productivity of the last supplier lost in ROW rises, while the productivity of the last supplier gained in *Foreign* falls.

A larger *Home* (higher β) shifts each point of the downward-sloping function upwards, yielding higher $\widetilde{\omega}_N$, $\widetilde{\omega}_{ROW}$ and $\widetilde{\omega}_F$. Intuitively, with more buyers, the productivity of the marginal supplier falls in all jurisdictions with and without a PTA.

Now consider the effect of *Foreign* becoming small relative to ROW. This is represented by a fall in γ. In that case, $\widetilde{\omega}_N$ does not change, because the cutoffs under no PTA do not depend on the relative size of *Foreign*. But the cutoffs under the PTA do change. The downward-sloping function pivots around the $\widehat{\omega}_F = \widehat{\omega}_{ROW} = \widetilde{\omega}_N$ point and becomes steeper, while the y-axis intercept $F^{-1}\left(\frac{\beta}{\gamma}\right)$ rises. The cutoffs $\widetilde{\omega}_F$ and $\widetilde{\omega}_{ROW}$ both rise.[18] However, note that the decrease in the cutoff in ROW induced by the PTA, $\widetilde{\omega}_N - \widetilde{\omega}_{ROW}$, becomes smaller as γ falls, while the counterpart increase in the cutoff in *Foreign* induced by the PTA, $\widetilde{\omega}_F - \widetilde{\omega}_N$, gets larger as γ falls. Intuitively, under a lower γ suppliers in *Foreign* become relatively more scarce, so the PTA induces suppliers lower down in the productivity distribution to form vertical chains with buyers.

---

[18]Mathematically, the effect of a higher γ is $\frac{d\widetilde{\omega}_{ROW}}{d\gamma} = \frac{F(\widetilde{\omega}_{ROW}) - F(\widetilde{\omega}_{ROW}+t)}{\gamma f(\widetilde{\omega}_{ROW}+t) + (1-\gamma)f(\widetilde{\omega}_{ROW})} < 0.$

5   The Welfare Consequences of a PTA

We can express the welfare generated by a single Y-chain without a trade agreement and under a PTA as, respectively,[19]

$$\Psi_N(\omega) = [V(Q^*) - p_w Q^*] + p_w q_N^* - C(q_N^*, i_N^*) - I(i_N^*) \text{ and} \tag{23}$$

$$\Psi_P(\omega, t) = [V(Q^*) - p_w Q^*] + p_w q_P^* - C(q_P^*, i_P^*) - I(i_P^*). \tag{24}$$

The first bracketed term is identical in the two expressions and reflects the fact that, by design, consumer welfare from the final good remains constant regardless of whether a PTA obtains. Hence, the PTA has no effect on it. The other terms of $\Psi_i(\omega)$ denote the surplus—including government's tariff revenue—created when a vertical chain forms under trade regime i, relative to the surplus B would generate if he only bought generic inputs from ROW. Observe that, in the limiting case where the tariff is very small, $\lim_{t \to 0} \Psi_P = \Psi_N$. We denote the welfare impact of the PTA due to a single Y-chain where the supplier has parameter ω by $\Delta\Psi(\omega, t) \equiv \Psi_P(\omega, t) - \Psi_N(\omega)$.

We obtain the total welfare impact of a PTA by aggregating the effects over all Y-chains. Welfare without trade agreements is given by

$$W_N = \int_0^{\widetilde{\omega}_N} \Psi_N(\omega) dF(\omega),$$

while welfare under a PTA satisfies

$$W_P = \gamma \int_0^{\widetilde{\omega}_F(t)} \Psi_P(\omega, t) dF(\omega) + (1 - \gamma) \int_0^{\widetilde{\omega}_{ROW}(t)} \Psi_N(\omega) dF(\omega).$$

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

$$\Delta W(\gamma) = \underbrace{\gamma \int_0^{\widetilde{\omega}_N} \Delta\Psi(\omega, t) dF(\omega)}_{\textit{incumbent supplier effect: } IS(\gamma)} + \underbrace{\left[\gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} \Psi_P(\omega, t) dF(\omega) - (1 - \gamma) \int_{\widetilde{\omega}_{ROW}(t)}^{\widetilde{\omega}_N} \Psi_N(\omega) dF(\omega) \right]}_{\textit{new supplier effect: } NS(\gamma)}. \tag{25}$$

The first term of (25) corresponds to the welfare impact of the PTA for all Y-chains with specialized suppliers in *Foreign,* and where the $B - S$ vertical chain forms both with and without

---

[19]In the Appendix we show these expressions under the functional forms we adopted.

the PTA. We refer to this as the aggregate *incumbent supplier effect*, and denote it by $IS(\gamma)$. The term in brackets corresponds to the welfare impact due to the reallocation of buyers from vertical chains with suppliers from ROW (outside the PTA) to vertical chains with suppliers from *Foreign* (inside the PTA). We refer to this as the aggregate *new supplier effect*, and denote it by $NS(\gamma)$. We can then write $\Delta W(\gamma) = IS(\gamma) + NS(\gamma)$.

For expositional reasons, it is best to investigate expression (25) in parts. In subsection 5.1 we analyze the welfare consequences of a PTA for an incumbent supplier in *Foreign* where the supplier's productivity ω is arbitrary. From subsection 5.2 onwards we then consider the aggregate welfare impact of the PTA across all ω, taking into account changes in the set of Y-chains. However, to distinguish across various forces, we first consider the case where $\gamma = 1$. In that case, there are no new vertical chains, so $NS(1) = 0$ and $\Delta W(1) = IS(1)$. We can think of that as the limiting situation of cases where the preferential partner is very large, e.g., the US for Mexico within NAFTA. Or more generally, it can represent (the extreme version of) cases where the PTA members are strong "natural partners," perhaps due to geographical remoteness, as for example Australia and New Zealand. Analytically, setting $\gamma = 1$ allows us to keep the set of vertical chains unchanged by the PTA. In subsection 5.3 we focus instead on the "extensive margin" effects of the PTA, highlighting how changes in the set of vertical chains due to the PTA influences its total welfare impact. That is, we analyze $NS(\gamma)$ in isolation. Finally, in subsection 5.4 we analyze $\Delta W(\gamma)$ for general γ.

5.1 Single Partnership

Within a given incumbent vertical chain, 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. It is instructive to split $\Delta \Psi (\omega, t)$ into two effects, *relationship strengthening* $(\Delta \Psi_R)$ and *sourcing diversion* $(\Delta \Psi_S)$, with $\Delta \Psi (\omega, t) = \Delta \Psi_R + \Delta \Psi_S$.

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 from 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,

$$\Delta \Psi_R = p_w(q_1^* - q_N^*) + [C(q_N^*, i_N^*) - C(q_1^*, i_P^*)] - [I(i_P^*) - I(i_N^*)]. \tag{26}$$

After some manipulation, this expression can be rewritten as

$$\Delta \Psi_R = \frac{2c - b^2}{2c} \Delta i \left(HUP_N - EXC_P \right). \tag{27}$$

Expression (27) is very intuitive. There is underinvestment in the absence of trade agreements $(HUP_N > 0)$, and the increase in investment $(\Delta i > 0)$ mitigates that original inefficiency. The first term in 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 $EXC_P > 0$. The second term in parenthesis reflects the welfare losses from inducing the supplier to invest above the first-best level. The sign of $\Delta \Psi_R$ depends upon which of those two gaps is more egregious. Naturally, if the underinvestment problem remains present under the PTA despite the extra investment, then $EXC_P < 0$ and $\Delta \Psi_R > 0$ for sure.

It also follows from expression (27) that $\Delta \Psi_R$ is non-monotonic in Δi. When Δi is small, the relationship-strengthening effect is positive and increasing in Δi. But when Δi is very high, $HUP_N - EXC_P < 0$ and an increase in Δi amplifies the distortion in investment spending.

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 customized 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,

$$
\begin{aligned}
\Delta \Psi_S &= C(q_1^*, i_P^*) - C(q_P^*, i_P^*) + p_w(q_P^* - q_1^*) \\
&= -\frac{t^2}{2c}.
\end{aligned}
\tag{28}
$$

This corresponds to (the negative of) area $\mathbf{E}$ in Figure 2—a triangle with base $(q_P^* - q_1^*) = \frac{t}{c}$ and height t.

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., $\Delta \Psi_R \geq |\Delta \Psi_S|$. This compari-

son highlights a tradeoff between improvements in the efficiency of the production process ($\Delta \Psi_R$) versus tariff-induced allocative inefficiency ($\Delta \Psi_S$).

A key determinant of the balance of this tradeoff is the supplier's (inverse) productivity parameter, ω, which shifts her marginal cost function. From Lemma 1 we have that $\frac{\partial HUP_N}{\partial \omega} < 0$. And it is straightforward to see that $\left| \frac{\partial EXC_P}{\partial \omega} \right| = \left| \frac{\partial HUP_N}{\partial \omega} \right|$. It follows that productivity has a higher impact on the efficient level of investment than on the privately chosen level of i at any trade regime. Therefore, taking the partial derivative of (27), we find

$$\frac{\partial \Delta \Psi_R}{\partial \omega} = \frac{2c - b^2}{c} \Delta i \frac{\partial HUP_N}{\partial \omega} < 0. \tag{29}$$

This implies that the potential efficiency-enhancing aspect of a PTA is unambiguously more important for more productive suppliers (which have a lower ω). The key force behind this result is that the inefficiency brought about by contractual incompleteness is increasing in productivity. Thus, when cost-reducing investment rises with the PTA, it brings a greater welfare benefit for low-ω suppliers.

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 Y-chain, the downside of a PTA is unaffected by the productivity of the specialized supplier, whereas the upside rises with it. Thus, we have that:

Lemma 2 *Higher supplier productivity induces a stronger relationship-strengthening effect, but has no impact on the sourcing-diversion effect of a Preferential Trade Agreement. Hence, $\Delta \Psi(\omega, t)$ is decreasing in ω.*

A central element behind Lemma 2 is that only the slope (and not the level) of the marginal cost curve affects the sourcing-diversion effect. Since productivity only shifts that curve vertically, productivity does not influence the extent of sourcing diversion.

An implication of Lemma 2 is that, considering a single partnership, the PTA raises welfare ($\Delta \Psi_R + \Delta \Psi_S \geq 0$) if

$$\omega \leq p_w - \left[\frac{2c - 2\alpha b^2 + \alpha^2 b^2}{2\alpha(1 - \alpha)b^2} \right] t \equiv \bar{\omega}. \tag{30}$$

Observe that, since $2c > b^2$, the expression in brackets is strictly positive. Furthermore, note that $\bar{\omega}$ is negative if t is sufficiently high. In that case, there are no Y-chains for which the PTA enhances welfare. We therefore have that:

Lemma 3 *If*

$$t > \left[\frac{2\alpha(1-\alpha)b^2}{2c - 2\alpha b^2 + \alpha^2 b^2} \right] p_w \equiv \bar{t}, \tag{31}$$

then the PTA lowers welfare for all existing Y-chains.

Proof. If condition (31) holds, $\bar{\omega} < 0$. Hence, the PTA lowers welfare for all existing Y-chains. ∎

Lemma 3 places some bounds on the benefits of a PTA stemming from the relationship-strengthening effect. Specifically, the PTA is unable to raise welfare due to a given vertical chain if the margin of preference is too high. Similarly, 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 placing tight bounds on $\bar{t}$. An analogous point can be made for very low levels of b.

On the other hand, if the external tariff is sufficiently small, then the net within-chain impact of a PTA is necessarily positive. See the Appendix for the proof.

Lemma 4 *The within-chain impact of a PTA is positive when the external tariff is very small:* $\frac{d\Delta\Psi(\omega,t)}{dt}(t=0) > 0.$

Hence, if the external tariff is sufficiently small, the first-order gain from the relationship-strengthening effect dominates the second-order loss from the sourcing-diversion effect within an existing Y-chain.

5.2 Aggregate Welfare Impact when *Foreign* is Large ($\gamma = 1$)

When $\gamma = 1$, $\widetilde{\omega}_{ROW} = \widetilde{\omega}_F = \widetilde{\omega}_N$. The PTA affects only vertical chains where *Foreign* suppliers are matched with or without the PTA. The entire welfare effect is due to those incumbent vertical chains, and equation (25) reduces to simply

$$\Delta W(1) = IS(1) = \int_0^{\widetilde{\omega}_N} \Delta\Psi(\omega,t)dF(\omega). \tag{32}$$

27

In that case, the PTA affects welfare only through the relationship-strengthening and the sourcing-diversion effects, aggregated over all existing Y-chains. We term the aggregate effects due to those two forces RS and SD, respectively. We know from the previous analysis that $SD < 0$ but that in general the sign of RS is ambiguous.

Now, since Lemma 2 shows that $\Delta\Psi(\omega, t)$ decreases with ω, it follows immediately that, if $\widetilde{\omega}_N \leq \bar{\omega}$, then $\Delta W(1) > 0$. That is, if the PTA is not harmful even through the marginal active vertical chain, then it is overall helpful for sure. In that case, the distribution of active suppliers is restricted to those for which the welfare impact of the PTA is positive. If instead $\widetilde{\omega}_N > \bar{\omega}$, then whether the PTA helps or hurts in aggregate hinges on the whole distribution of productivity of the active specialized suppliers.

Because of Lemma 2 we can, however, rank distributions. In particular, let us say that $F_2(\omega)$ **FOSD** $F_1(\omega)$ when distribution $F_2(\omega)$ first-order stochastically dominates distribution $F_1(\omega)$. In that case, we have that a PTA yields better welfare consequences under $F_1(\omega)$ than under $F_2(\omega)$. See the Appendix for the proof.

Proposition 1 *If $F_2(\omega)$ **FOSD** $F_1(\omega)$, then $\Delta W(1; F_1) > \Delta W(1; F_2)$.*

Proposition 1 implies that, in the context of global sourcing, a PTA enhances welfare provided that the distribution of active specialized suppliers is sufficiently concentrated on high-productivity suppliers, but not otherwise. A corollary is that, if one were able to identify a distribution $F_0(\omega)$ under which a PTA would be welfare-neutral, one would know that the agreement would be socially desirable under all distributions that are "better" than $F_0(\omega)$, in the sense of being first-order stochastically dominated by $F_0(\omega)$, and undesirable under all distributions with the opposite property.

Proposition 1 could also 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 j such that $F_j(\omega)$ **FOSD** $F_0(\omega)$ should be excluded from the agreement, whereas all industries i such that $F_0(\omega)$ **FOSD** $F_i(\omega)$ should be integral parts of it.

Now, a central element determining the social desirability of a PTA is the level of the external tariff, which defines the extent of preferential treatment for matches in *Foreign*. It affects RS and

SD differently.[20]

While in general the effect of a higher external tariff is ambiguous, we know what happens at the extremes. Lemma 3 states that, if t is too high, then a PTA lowers welfare through all existing Y-chains and is therefore definitely harmful. Yet if the external tariff is sufficiently small, then it follows from Lemma 4 that a PTA raises welfare through all existing Y-chains and is therefore surely beneficial. Indeed, we will see that the effect of t on $\Delta W(1)$ is non-monotonic.

On one hand, sourcing diversion is a very simple function of the external tariff, monotonically increasing with t at an increasing rate. On the other hand, the relationship-strengthening effect is more nuanced. For a given Y-chain, it is positive for sufficiently low t, initially rises, but eventually falls with t.

For very low t, SD is second-order small, so RS dominates. But because the tariff is small, the change in investment is also small, and so is RS. Thus, the effects of the PTA are minor. As t increases, Δi increases. For relatively low levels of t, the welfare gain from a PTA rises with t. For sufficiently high t, however, the increase in RS is more than offset by a fall in SD, and the welfare gain from a PTA falls with the external tariff. Thus, for any distribution of ω, there is a maximum level of t that is consistent with welfare-improving PTAs. See the Appendix for the proof.

Proposition 2 *When $\gamma = 1$, the welfare impact of a PTA has an inverted-U shape with respect to the external tariff. It is strictly positive when the external tariff is sufficiently close to zero, is maximized when $t = \hat{t}$, where $\hat{t}$ corresponds to*

$$\hat{t} \equiv \frac{\alpha(1-\alpha)b^2\left[p_w - \mathbf{E}\left(\omega; \omega \leq \widetilde{\omega}_N\right)\right]}{2c - 2\alpha b^2 + \alpha^2 b^2}, \tag{33}$$

and is strictly negative when $t > 2\hat{t}$.

Hence, there is a level of preferential margin $\hat{t}$ that optimally trades off the gains from RS against the losses from SD.[21] The same factors that determine $\hat{t}$ also determine the highest level of preferential margin under which a PTA can be beneficial, which here is simply $2\hat{t}$. Both are

---

[20]Naturally, the tariff also affects welfare through the conventional mechanism of inefficiently lowering the total volume of traded inputs, Q^*. However, under dual sourcing with and without the PTA, that effect is unchanged by the agreement.

[21]Observe that $\mathbf{E}\left(\omega; \omega \leq \widetilde{\omega}_N\right)$ is fully determined by the distribution of ω and by parameter β, so $\hat{t}$ is a function of primitives only.
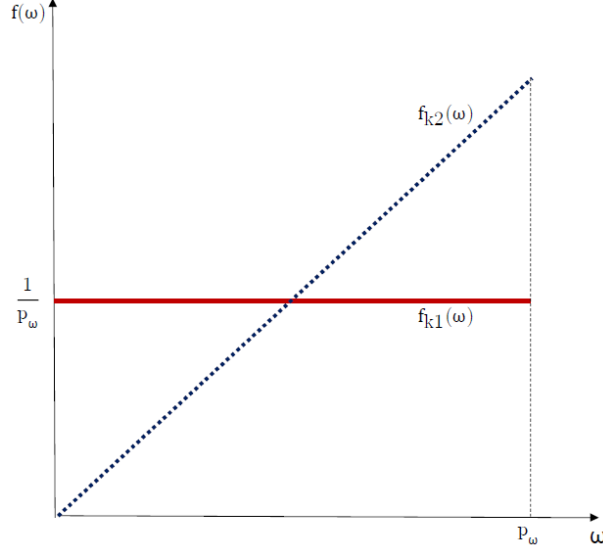
Fig. 4: Densities for $k = 1$ and $k = 2$

an increasing function of the average productivity of the active specialized suppliers [i.e., $\hat{t}$ rises as $\mathbf{E}(\omega; \omega \leq \widetilde{\omega}_N)$ falls]. This happens because, 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 higher margin of preference to boost RS. It is also intuitive that a higher b generates a greater $\hat{t}$, since b represents the sensitivity of marginal cost to investment, which is boosted by the external tariff.

Example 1 *To illustrate both propositions, consider that fundamental productivity $1/\omega$ follows a Pareto distribution with lower distribution bound $1/p_w$ and shape parameter $k \geq 1$. This yields $F(\omega) = \left(\frac{\omega}{p_w}\right)^k$ for $\omega \in [0, p_w]$. Consider then the distributions for $k = 1, 2$, $F_{k1}(\omega) = \frac{\omega}{p_w}$ and $F_{k2}(\omega) = \left(\frac{\omega}{p_w}\right)^2$. $F_{k1}(\omega)$ corresponds to a uniform distribution. It is obvious that $F_{k2}(\omega)$ **FOSD** $F_{k1}(\omega)$. Equilibrium cutoffs are $\widetilde{\omega}_{N1} = \beta p_w$ and $\widetilde{\omega}_{N2} = \sqrt{\beta} p_w$, and $\mathbf{E}(\omega; \omega \leq \widetilde{\omega}_{N1}) < \mathbf{E}(\omega; \omega \leq \widetilde{\omega}_{N2})$. Figure 4 shows the two densities, while Figure 5 shows $\Delta W(1)$ for each of them as a function of the tariff.[22] Following Proposition 1, $\Delta W(1)$ is higher for every t under $F_{k1}(\omega)$. Following Proposition 2, for both distributions $\Delta W(1)$ is an inverted-U with respect to t, is strictly positive for small external tariffs, and is strictly negative for tariffs more than twice as large the tariff that maximizes it. Furthermore, the peak of $\Delta W(1)$ obtains for a higher t under $F_{k1}(\omega)$.*

---

[22] Figure 5 assumes $p_w = c = 1$, $b = 1.25$ and $\alpha = \beta = 0.5$. There is nothing special about this parametrization.
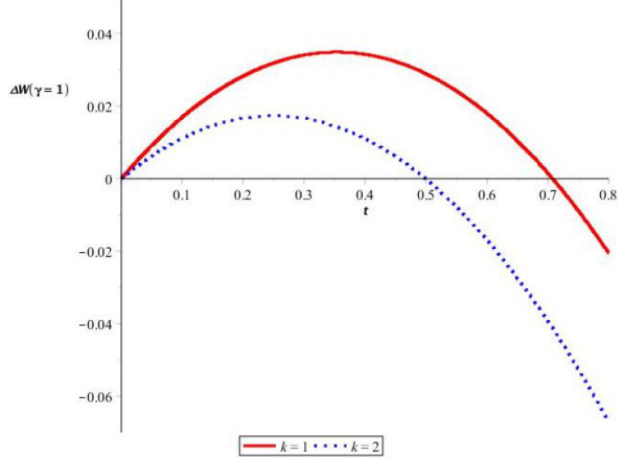
Fig. 5: $\Delta W(\gamma = 1)$ for $k = 1$ and $k = 2$

5.3 The New Supplier Effect

In the previous subsection we analyzed in detail the incumbent supplier effect of a PTA when $\gamma = 1$. The general $IS(\gamma)$ is simply $\gamma IS(1)$. Hence, if $\gamma < 1$ the analysis of that term remains the same, but the welfare impact is of lower magnitude. The remaining part of the welfare impact is the new supplier effect, $NS(\gamma)$, to which we turn now.

The new supplier effect is defined as

$$NS(\gamma) \equiv \gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} \Psi_P(\omega, t) dF(\omega) - (1 - \gamma) \int_{\widetilde{\omega}_{ROW}(t)}^{\widetilde{\omega}_N} \Psi_N(\omega) dF(\omega). \tag{34}$$

The first term measures welfare generated by suppliers in *Foreign* that join vertical chains under the PTA but not without a PTA. The second term measures welfare generated by "old" suppliers in *ROW* that join vertical chains under no PTA but do not join vertical chains under a PTA. This term enters negatively because those suppliers are effectively replaced after the agreement. The new supplier effect is complicated because there is both a change in the distribution of supplier productivity and a set of new investment levels due to the tariff preference under the PTA. The productivity cutoffs $\widetilde{\omega}_{ROW}(t)$ and $\widetilde{\omega}_F(t)$ are different from the old cutoff $\widetilde{\omega}_N$ that obtains in *ROW* and *Foreign* under no PTA, and welfare $\Psi_P(t, \omega)$ depends upon the new levels of investment.

31

To simplify the analysis, it is useful to express $NS(\gamma)$ in a slightly different form:

$$NS(\gamma) \equiv \gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} \Delta\Psi(\omega, t)dF(\omega)$$
$$+ \left[\gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} \Psi_N(\omega, t)dF(\omega) - (1-\gamma) \int_{\widetilde{\omega}_{ROW}(t)}^{\widetilde{\omega}_N} \Psi_N(\omega)dF(\omega) \right]. \quad (35)$$

The first term of (35) is similar to $IS(\gamma)$, except that it covers vertical chains with supplier productivity $\omega \in (\widetilde{\omega}_N, \widetilde{\omega}_F]$ instead of $\omega \in [0, \widetilde{\omega}_N]$. The second (bracketed) term is fundamentally different. It represents the welfare consequences of the PTA due to the changes in the set of vertical chains, stripped from the within-Y-chain changes induced by the elimination of tariffs on imports from *Foreign*. We term it the *matching diversion effect*, and denote it as $MD(\gamma)$. The following result shows that it is always negative. See the Appendix for the proof.

Proposition 3 *For any $t > 0$, the matching diversion effect due to a PTA is negative.*

Because of the tariff preference, some buyers with less-than-great matches in ROW rematch in *Foreign*. The new vertical chains include worse suppliers than the original ones. Hence, if we disregard the changes in investment and production due to the tariff preferences, this inefficient reallocation of suppliers across vertical chains necessarily lowers global welfare.

Now, the tariff preference could induce socially beneficial changes in investment and production that outweigh the matching diversion effect, as we illustrate later in this subsection. But it turns out 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. For ease of exposition, we first identify two sufficient conditions for $NS(\gamma) < 0$, one on the tariff and another on the density. We then identify the pair of conditions necessary for $NS(\gamma) > 0$, and introduce an example highlighting them.

Consider the tariff. If it is too high, then changes in investment fail to yield a positive welfare effect for the cutoff supplier under no PTA, $\widetilde{\omega}_N$. Specifically, for any tariff high enough so that $\Delta\Psi(\widetilde{\omega}_N, t) \leq 0$, Lemma 2 implies that all "new" suppliers ($\omega \in (\widetilde{\omega}_N, \widetilde{\omega}_F(t)]$) generate lower welfare under the PTA and the first term in (35) is surely negative. It follows that the whole new supplier effect must be negative in that case. See the Appendix for the proof.

Proposition 4 *If $t \geq \frac{2\alpha(1-\alpha)b^2 \left[p_w - F^{-1}(\beta) \right]}{2c - 2\alpha b^2 + \alpha^2 b^2} \equiv t^{NS}$, then the new supplier effect is negative.*

If $t < t^{NS}$, then $\Delta\Psi(\widetilde{\omega}_N, t) > 0$ and the first term in (35) may be positive. But this is by no means sufficient for $NS(\gamma) > 0$. Indeed, for certain densities of suppliers, the matching diversion effect dominates for *any* t.

To analyze the role played by the density, it is helpful to delve deeper into the mechanics of supplier reallocation. Intuitively, for tariff t, there is a reallocation of buyers from vertical chains with ROW suppliers ($\omega \in [\widetilde{\omega}_{ROW}(t), \widetilde{\omega}_N]$) to vertical chains with *Foreign* suppliers ($\omega \in [\widetilde{\omega}_N, \widetilde{\omega}_F(t)]$). For a small change in the tariff from t to $t + dt$, the cutoff supplier $\widetilde{\omega}_{ROW}(t)$ falls, the cutoff supplier $\widetilde{\omega}_F(t)$ rises, and an additional number of supplier reallocations occur. The exact measure of reallocations induced by the increase dt is a function of both the density of cutoff suppliers in ROW, $\gamma f(\widetilde{\omega}_{ROW}(t))$, and the density of cutoff suppliers in *Foreign*, $(1 - \gamma)f(\widetilde{\omega}_F(t))$.

To make it easy to think about this measure, we call it the *flow rate* of reallocations. We can derive a precise expression for this flow rate by using a change of variables to rewrite (34) as:[23]

$$NS(\gamma) = \int_0^t \left[\Psi_P(\widetilde{\omega}_F(x), t) - \Psi_N(\widetilde{\omega}_{ROW}(x)) \right] \phi(x; \gamma, F) dx. \qquad (36)$$

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

$$r(x, t) \equiv \Psi_P(\widetilde{\omega}_F(x), t) - \Psi_N(\widetilde{\omega}_{ROW}(x)).$$

The reallocation 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 $\widetilde{\omega}_{ROW}(x)$ in ROW and forms a new one with supplier $\widetilde{\omega}_F(x)$ in *Foreign*, but invests and produces according to external tariff t.

In turn, the function $\phi(x; \gamma, F)$ captures precisely the flow rate of buyers that (due to the PTA) move from vertical chains with ROW suppliers with productivity $\widetilde{\omega}_{ROW}(x)$ to new vertical chains

---
[23]See the Appendix for the derivation of this expression.

with *Foreign* suppliers with productivity $\widetilde{\omega}_F(x)$. Specifically, we have

$$\phi(x; \gamma, F) \equiv \frac{\gamma(1-\gamma)f(\widetilde{\omega}_F(x))f(\widetilde{\omega}_{ROW}(x))}{\gamma f(\widetilde{\omega}_F(x)) + (1-\gamma)f(\widetilde{\omega}_{ROW}(x))}.$$

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 total effect $NS(\gamma)$ aggregates the reallocation function over all supplier reallocations that occur under the PTA according to the weights $\phi(x; \gamma, F)$. We now state a monotonicity condition.

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

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

$$(1-\gamma)f(\widetilde{\omega}_{ROW}(x))^3 f'(\widetilde{\omega}_F(x)) - \gamma f(\widetilde{\omega}_F(x))^3 f'(\widetilde{\omega}_{ROW}(x)) \geq 0.$$

With a uniform distribution, $f_{k1}(\omega) = \frac{1}{p_w}$, the flow rate of new reallocations is constant and satisfies Condition 1 for any γ and t. The condition is restrictive, however. For other distributions, such as $f_{k2}(\omega) = \frac{2\omega}{p_w^2}$, it is often the case that it holds for some γ and t, but not all. Still, if Condition 1 holds, then $NS(\gamma) < 0$ regardless of t. See the Appendix for the formal proof.

Proposition 5 *Under Condition 1, $NS(\gamma) < 0$ for any positive t.*

Intuitively, if the flow rate of reallocations rises with the size of the tariff, then there are relatively more reallocations at the margin than inframarginally. As a result, any welfare improvement from higher investments is dominated by welfare losses due to matching diversion.

We provide here a sketch of the proof, which rests on two observations: (1) For $t = 0$, $NS(\gamma) = 0$; and (2) under Condition 1, $NS(\gamma)$ is decreasing and concave. The first observation is obvious, so let t be positive. For relatively efficient reallocations, x is near 0. At $x = 0$, the welfare effect $r(0, t) = \Psi_P(\widetilde{\omega}_N, t) - \Psi_N(\widetilde{\omega}_N) = \Delta\Psi(\widetilde{\omega}_N)$ is the same as the welfare impact of the PTA due to the marginal no-PTA supplier $\widetilde{\omega}_N$, and may be positive or negative. But as x increases, $r(x, t)$ unambiguously falls.

Lemma 5 *The reallocation function is decreasing in x.*

Proof. Differentiating, we have

$$\frac{dr(x,t)}{dx} = \frac{d\Psi_P(\widetilde{\omega}_F(x,\gamma),t)}{d\widetilde{\omega}_F}\frac{d\widetilde{\omega}_F}{dx} - \frac{\Psi_N(\widetilde{\omega}_{ROW}(x,\gamma))}{d\widetilde{\omega}_{ROW}}\frac{d\widetilde{\omega}_{ROW}}{dx},$$

which is negative because $\frac{d\Psi_P(\widetilde{\omega}_F(x),t)}{d\widetilde{\omega}_F} < 0$, $\frac{d\widetilde{\omega}_F}{dx} > 0$, $\frac{\Psi_N(\widetilde{\omega}_{ROW}(x))}{d\widetilde{\omega}_{ROW}} < 0$ and $\frac{d\widetilde{\omega}_{ROW}}{dx} < 0$. ∎

Intuitively, as x increases, the productivity of the old ROW supplier $\widetilde{\omega}_{ROW}(x)$ improves and the productivity of the new *Foreign* supplier $\widetilde{\omega}_F(x)$ worsens. Hence, the productivity gap between old and new suppliers grows with x. This lowers the welfare effect of reallocation 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 2 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. At that point, the net joint profits generated with supplier $\widetilde{\omega}_F(t)$ in *Foreign* under the PTA and with supplier $\widetilde{\omega}_{ROW}(t)$ in ROW without the PTA are the same. 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: $-tq_P^*(\widetilde{\omega}_F(t))$. Hence, the matching process induces welfare losses for sure at the margin, even after accounting for potentially beneficial changes in investment.

Now, if the flow rate of new matches with productivity near $\widetilde{\omega}_N$ is the same as the flow rate of new matches with suppliers with productivity near $\widetilde{\omega}_F$, then the negative effects due to the latter group of rematches will dominate and make $NS(\gamma) < 0$. 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 then follows that $NS(\gamma)$ is decreasing and concave in t.

Figure 6 illustrates the reallocation function and its relationship to $\Delta\Psi(\omega)$. For this comparison, it is helpful to change variables in the r function once more. We can write

$$NS(\gamma) = \gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} r(\omega,t)dF(\omega),$$
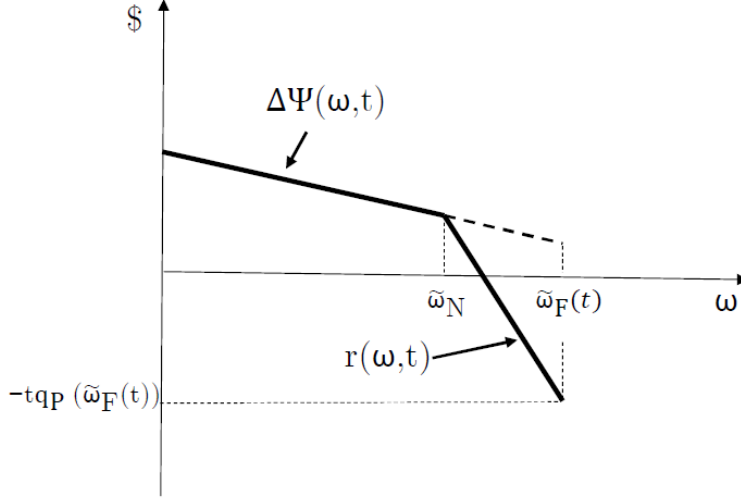
Fig. 6: Welfare Effects and the Reallocation Function

where

$$r(\omega, t) \equiv \Psi_P(\omega, t) - \Psi_N(\widetilde{\omega}_{ROW}(\omega))$$

shows, for an arbitrary external tariff t, the welfare impact of the PTA due to each reallocation to ω in *Foreign* from $\widetilde{\omega}_{ROW}(\omega)$ in *ROW*. For $\omega \le \widetilde{\omega}_N$, $\Delta\Psi(\omega, t)$ denotes the impact due to each incumbent supplier in *Foreign*. Because of Lemma 2, this function is decreasing in ω. The whole $IS(\gamma)$ aggregates over $\Delta\Psi(\omega, t)$ from 0 to $\widetilde{\omega}_N$ according to the density $f(\omega)$.

The dashed line is the welfare impact that the PTA would have for suppliers distributed over $(\widetilde{\omega}_N, \widetilde{\omega}_F]$ *if they were incumbent.* But they are not. Instead, they replace suppliers distributed over $[\widetilde{\omega}_{ROW}, \widetilde{\omega}_N)$ previously matched in *ROW*. The difference between the dashed line and the solid line to the right of $\widetilde{\omega}_N$ represents the loss due to matching diversion. This effect is negligible for the very first rematches, but grows large as reallocation continues. As t rises, the $r(\omega, t)$ portion of the curve necessarily lengthens, since $\widetilde{\omega}_F$ increases with t. The whole $NS(\gamma)$ aggregates over $r(\omega, t)$ from $\widetilde{\omega}_N$ to $\widetilde{\omega}_F$.

Unlike our analysis of $IS(\gamma)$, it is not straightforward to use first-order stochastic dominance to rank new supplier effects. The reason is that the "worse" distribution of productivity could have a low-magnitude new supplier effect [if the density is very low between $\widetilde{\omega}_{ROW}(t)$ and $\widetilde{\omega}_F(t)$], while the "better" distribution could have a severely negative new supplier effect [if the density happens to be very high around $\widetilde{\omega}_{ROW}(t)$ and $\widetilde{\omega}_F(t)$]. We can still make inferences, though. For example,

36

in comparing new supplier effects $NS_2(\gamma)$ and $NS_1(\gamma)$ for distributions where all that is known is that $F_2(\omega)$ **FOSD** $F_1(\omega)$, we could have that $NS_2(\gamma) < 0$ for all $F_2(\omega)$ densities while $NS_1(\gamma) > 0$ for some $F_1(\omega)$. But the opposite would be impossible.

Observe that, in the example displayed in Figure 6, $\Delta\Psi(\widetilde{\omega}_N) > 0$. This implies that every incumbent supplier contributes more to social welfare under the PTA than otherwise. It also implies that $NS(\gamma)$ *can* be positive. Propositions 4 and 5 imply the following necessary condition.

Corollary 1 *The new supplier effect is positive only if $t < t^{NS}$ and the flow rate $\phi(x;\gamma,F)$ is strictly decreasing for some x.*

Intuitively, if Condition 1 fails to hold, then $NS(\gamma)$ may be convex in t for some range of t and can be positive.[24] The next example illustrates that, for a given set of parameters and for a given tariff (below t^{NS}), we can always construct a density such that $NS(\gamma)$ is positive.

Example 2 *Let $t < t^{NS}$, $\gamma = \frac{1}{2}$ and*

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

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

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

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

$$
NS\left(\frac{1}{2}\right) = \frac{1}{2}\left[\eta \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_N+\widehat{\varepsilon}} r(\omega,t)d\omega + \left(\frac{1-2\widehat{\varepsilon}\eta}{1-2\widehat{\varepsilon}}\right)\int_{\widetilde{\omega}_N+\widehat{\varepsilon}}^{\widetilde{\omega}_F} r(\omega,t)d\omega\right].
$$

---

[24]Condition 1 addresses one of many terms in the second derivative of $NS(\gamma)$ with respect to t. It is frequently the case that other terms overwhelm the effects of a decreasing flow rate. For example, $NS(\gamma) < 0$ and is strictly concave under the Pareto ($k = 2$) distribution of Example 1, even though it does not (always) satisfy Condition 1.
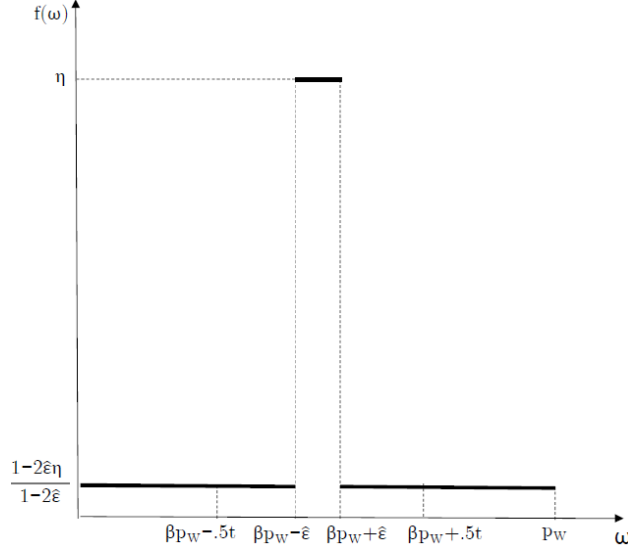
Fig. 7: A Distribution that Yields a Positive New Supplier Effect

It follows that $\int_{\widetilde{\omega}_N}^{\widetilde{\omega}_N+\widehat{\varepsilon}} r(\omega, t)d\omega > 0$ and $\int_{\widetilde{\omega}_N+\widehat{\varepsilon}}^{\widetilde{\omega}_F} r(\omega, t)d\omega < 0$. Hence, for η sufficiently close to $\frac{1}{2\widehat{\varepsilon}}$, the new supplier effect is positive. Figure 7 highlights the intuition. If the density of idle suppliers (under no PTA) in Foreign *is very high for supplier reallocations very close to $\widetilde{\omega}_N$, and this density is very low for other supplier reallocations, then it is possible to have a positive new supplier effect. Compare Figure 7 with Figure 4. The density $f_{PU}(\omega)$ distorts $f_{k1}(\omega)$, allocating more density near $\widetilde{\omega}_N$ and less density near $\widetilde{\omega}_{ROW}$ and $\widetilde{\omega}_F$. But it does not alter the equilibrium cutoffs $\widetilde{\omega}_{ROW}, \widetilde{\omega}_N$ and $\widetilde{\omega}_F$. Essentially, this reflects a situation where: (1)* Foreign *has a large number of suppliers with productivity near $\widetilde{\omega}_N$ 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 $\widetilde{\omega}_N$.*

Note that in this example, if $t > t^{NS}$, then no positive $\widehat{\varepsilon}$ exists and it is impossible to construct a density that yields $NS(\gamma) > 0$.

5.4 The General Case

We now consider the general case. The welfare consequences of the PTA comprise the sum of the aggregate incumbent supplier effect and the aggregate new supplier effect.

The sign of $IS(\gamma)$ 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.

38

The same analysis applies to the first component of $NS(\gamma)$ in equation (35) for the vertical chains that are formed in *Foreign* because of the PTA. Thus, its sign depends on the same forces that shape the first term. On the other hand, the second component of $NS(\gamma)$ in equation (35)—the matching diversion effect—is necessarily negative.

In general, then, a PTA in the context of global sourcing will raise aggregate welfare when incumbent supplier effects are sufficiently strong relative to any negative new supplier effects. While the net effect of those forces will in general be ambiguous—keeping up with the tradition of the regional integration literature—there are forces that tilt the balance in one direction or the other.

As discussed in the previous subsection, when $\gamma < 1$ the welfare effect of the PTA is not necessarily higher for a better distribution of productivity. When we consider the effects of tariffs on $\Delta W(\gamma)$, however, some of the results from the "large partner" ($\gamma = 1$) case go through. First, for a sufficiently low tariff, the total effect is unambiguously positive. Basically, the aggregate incumbent supplier effect is always positive for a sufficiently low t, while the aggregate new supplier effect is negligible for very low t. Second, the welfare effect of the PTA is negative if the tariff is sufficiently high. For the tariff such that $IS(\gamma) = 0$, $t = 2\hat{t}$, it is always true that if $\gamma < 1$, then $NS(\gamma) < 0$. Hence, the range of tariffs such that the PTA enhances welfare is smaller when $\gamma < 1$. See the Appendix for the proof.

Proposition 6 *For any $\gamma < 1$, there exists a $\underline{t} > 0$ such that if $t < \underline{t}$, then the PTA enhances aggregate welfare. Also, there exists a $\overline{t} \in [\underline{t}, 2\hat{t})$ such that if $t > \overline{t}$, then the PTA lowers aggregate welfare. Under Condition 1, $\underline{t} = \overline{t}$ is unique.*

An immediate implication of Proposition 6 is that, if the PTA in this context lowers aggregate welfare, it is because the external tariff—a policy variable that could potentially also be changed with the agreement—is too high.[25]

For $t \in (\underline{t}, \overline{t})$ when Condition 1 fails to hold, either of the aggregate effects may be positive or negative, but their signs are linked through the welfare effect of the PTA due to the marginal incumbent supplier. This is both the lowest possible welfare effect among incumbent suppliers, $\Delta\Psi(\widetilde{\omega}_N, t)$, and the highest possible reallocation effect, $r(0, t)$. If that term is positive, then the

---

[25]In fact, Crivelli (2016) shows empirically that external tariffs tend to fall upon the formation of free trade agreements especially when they are initially high.

welfare effect is positive for all incumbent suppliers and $IS(\gamma) > 0$. If it is negative, then the reallocation function is negative for all supplier reallocations and $NS(\gamma) < 0$. We can conclude that, if $IS(\gamma) < 0$, then we must have $\Delta\Psi(\widetilde{\omega}_N, t) < 0$. It then follows that $NS(\gamma) < 0$ and $\Delta W(\gamma) < 0$. On the other hand, if $IS(\gamma) > 0$, then it is possible that $\Delta\Psi(\widetilde{\omega}_N, t) > 0$, and $NS(\gamma)$ (as well as $\Delta W(\gamma)$) may be positive or negative.

Observe also that, under Condition 1, $NS(\gamma)$ is concave in t. Since $IS(\gamma)$ is also concave in t (Proposition 2), $\Delta W(\gamma)$ is as well, except that the external tariff that maximizes it is lower than $\hat{t}$.

Finally, the external tariff that maximizes welfare for a large PTA partner is inefficiently high for a smaller PTA partner. Intuitively, the tariff preference has a better effect when ω is lower. Thus, to maximize the aggregate incumbent supplier 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, $\Delta\Psi(\widetilde{\omega}_N, t) = r(0, t)$ is decreasing in t at $t = \hat{t}$ and welfare from all reallocations falls with t. We then have the following (see the Appendix for the proof).

Proposition 7 *If $\gamma < 1$, then $\Delta W(\gamma)$ is maximized for $t < \hat{t}$.*

6   Alternative Specifications for the Generic 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 impinge the greatest changes on our results is when *Foreign* is an exporter of g. 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 in that case for the sourcing of q from ROW, which would then be discriminated against g under the PTA. There would be, in particular, a relationship-*weakening* effect for vertical chains that are preserved in ROW after the PTA. The more general point of our analysis would nevertheless remain valid: a PTA improves the incentives to invest for the specialized suppliers whose inputs become relatively cheaper than the generic alternative, but worsens the incentives to invest for the specialized suppliers whose inputs

40

become relatively more expensive than the generic alternative because of the tariff preference.

For the reallocation of vertical chains, on the other hand, what matters is the relative tariff on specialized inputs from the two locations. Therefore, the essential insights from our previous analysis would remain unchanged if *Foreign* exported g, as they do not hinge on the location of the generic industry.

There are other possible specifications for the location of the generic industry, but they would have much less impact on our results. For example, if *Home* did not import but instead produced domestically all generic inputs that its buyers purchase, all of our novel results would remain unaltered. To see that, notice that all the action in the model hinges upon the difference between the tariffs applied to g and q. For notational simplicity, we defined the initial tariff to be the same, t, for both, but by now it should be clear that the whole analysis would carry through if the tariffs on g and q were, respectively, t_g and t_q, with $t_g \neq t_q$. Regarding the main insights of the analysis, the situation where *Home* produces all g it uses corresponds to the special case where $t_g = 0$.

Another possibility is when *Foreign* has an industry of generics but the industry is unable to supply enough g to fulfill *Home*'s demand, so *Home* still imports g 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 trade diversion in the sourcing of generic inputs, of the type analyzed by Grossman and Helpman (1995), which by now is well-known.

7   Deep Integration

A defining characteristic of all preferential trade agreements is the reduction of bilateral tariffs. However, PTAs are increasingly encompassing several other policies. These include the harmonization of product standards, bilateral recognition of intellectual property rights, rules providing investment protection, a common competition policy, etc.[26] Our framework can be readily extended to incorporate provisions like those. In fact, since such nontariff policies are likely to alter the effective level of investment protection for specialized suppliers, our framework is particularly well suited for that purpose.

---

[26]See, for example, World Trade Organization (2011) for a detailed discussion of the growing prevalence of those nontariff provisions in actual PTAs.

To analyze the differential impact of PTAs, let us then consider a simple extension of our benchmark model that incorporates bilateral recognition of intellectual property rights, focusing initially on a single partnership. Part of B's bargaining power could be due to its ability to sometimes costlessly copy S's technology. To capture this idea, suppose that after the investment is made but prior to bargaining over input production, nature determines whether S's technology is appropriable. With probability θ, the supplier's idea is not appropriable by the buyer, and they bargain over Ω_j, $j \in \{N, P\}$, as in the benchmark model. With probability $1 - \theta$, the buyer learns how to imitate and use S's technology to produce specialized inputs and the supplier earns zero revenue. The probability θ is a function of the stringency of bilateral recognition of IPRs.

The first-best level of investment remains the same. But the supplier's expected profit, net of the investment cost, is now $\theta\alpha\Omega_j$, and her problem under trade regime j is now

$$\max_i \theta\alpha\Omega_j - I(i).$$

Effectively, the supplier's bargaining power becomes $\alpha' \equiv \theta\alpha$. We term α' the level of *supplier investment protection*. The entire previous analysis carries through with α' replacing α.

Of course, myriad factors influence the determination of IPRs in an economy, but the modeling of how θ is determined is beyond the scope of this paper. We can, however, incorporate into our framework the possibility that a PTA brings about not only lower preferential tariffs but also provisions related to recognition of bilateral IPRs. A natural way to do so is to assume that a "deep PTA" both removes the tariff between *Home* and *Foreign and* puts in place rules/institutions that result in stricter recognition of bilateral intellectual property rights. Since such institutional changes may be difficult to alter, this is best modeled as a marginal increase in θ (and hence in α').[27]

Note, first, that $d\Delta q/d\theta > 0$, so deep integration is associated with a greater boost to bilateral trade flows, relative to "shallow integration" that only lowers tariffs. As indicated in the introduction, this is consistent with recent empirical findings. Furthermore, $d^2\Delta q/d\theta dt > 0$; thus, deep integration is complementary to shallow integration (i.e., a PTA that simply reduces bilateral tariffs) with respect to trade flows. Hence, the greater the tariff preference, the more effective deep

---

[27] Here we are following a modeling approach analogous to that of Osnago, Rocha and Ruta (2018), who model deep integration as an increase in the parameter governing contractibility, although they do so in the context of Antràs and Helpman's (2008) model.

integration is in terms of boosting bilateral trade. Entirely analogous statements can be made about the impact of a deep PTA on the investment effect, Δi.

Now, the *welfare* implications of deep integration are much more subtle. As the analysis of the previous section makes clear, the welfare impact of a shallow PTA already has several different components. To keep the analysis simple and to shed light on the effects of deep integration, we focus on the case when *Foreign* is a large, natural trading partner of *Home*; that is, when $\gamma = 1$.

To see how an increase in θ affects the welfare impact of a PTA, we need first to understand how supplier investment protection α' changes $\Delta W(1)$. Clearly, sourcing diversion effects are unaffected by α', but relationship-strengthening effects are, since α' determines the effective intensity of the hold-up problem. And recall that the agreement enhances overall welfare if it serves primarily to substitute for complete contracts for sufficiently productive firms, but not otherwise.

When investment protection is very strong (α' is near 1), there is no meaningful contractual inefficiency to substitute for. In that case, a PTA distorts sourcing decisions and induces excessive relationship-specific investment. In terms of equation (27), observe that when $\alpha' \to 1$, HUP_N vanishes but $EXC_P > 0$, so $RS < 0$ for any tariff. Thus, when $\alpha' \to 1$, the tariff discrimination under the PTA is necessarily harmful for society, as it generates sourcing diversion and a negative RS.

Conversely, when investment protection is seriously lacking (α' is near 0), the PTA is a poor substitute for contracts because the investment response to the PTA is too weak. In that case, the agreement merely distorts sourcing decisions. This is clear from (27), since $\lim_{\alpha' \to 0} \Delta i = 0$. Thus, also when $\alpha' \to 0$, the tariff discrimination under the PTA only brings undesirable effects.

In turn, when investment protection is neither too low nor too high, then PTAs can be effective. In that case, there is meaningful underinvestment but investment is sufficiently responsive to the tariff discrimination engendered by a PTA.

The next proposition formalizes those statements and shows how α' affects $\Delta W(1)$ more generally. See the Appendix for the proof.

Proposition 8 *When $\gamma = 1$, the welfare impact of the PTA is strictly negative when either $\alpha' \to 0$ or $\alpha' \to 1$, is increasing in α' when $\alpha' \to 0$ and decreasing in α' when $\alpha' \to 1$. Furthermore, it is*

maximized at an interior level α^O, defined as

$$\alpha^O \equiv \frac{2c \left[p_w - \mathbf{E} \left(\omega; \omega \leq \widetilde{\omega}_N \right) \right]}{\left(4c - b^2 \right) \left[p_w - \mathbf{E} \left(\omega; \omega \leq \widetilde{\omega}_N \right) \right] + \left(2c - b^2 \right) t}. \tag{37}$$

Hence, tariff preferences under a PTA cannot help if IPRs are too weak or the fundamental hold-up problem is too serious (as both lead to a very small α'), and cannot help either if IPRs are too strong and the fundamental hold-up problem is mild (as this would imply a very high α'). Instead, tariff preferences can help when both the original inefficiency and the stringency of IPRs are 'moderate.'

A direct consequence of Proposition 8 is that, when $\alpha' > \alpha^O$ an increase in α' through a higher θ *lowers* the welfare impact of the agreement, despite its positive effect on trade flows. The reason is that the beneficial role of the PTA in our context of international sourcing is to boost investment when investment is inefficiently low. When α' is already relatively high, further increasing it in the context of a PTA would bring little beneficial (and possibly excessive) investment coupled with sourcing diversion, thus decreasing the benefits of the agreement (and possibly turning them into a net loss).

On the other hand, when $\alpha' < \alpha^O$ a deep PTA has a higher welfare impact than a shallow agreement would. In that case hold-up problems are severe, and improving IPRs between the two PTA partners would boost the benefits brought about by the preferential tariff treatment, so there is a positive complementarity, from a social standpoint, between the effects of tariff discrimination and stricter IPRs on the supplier's investment. Thus, we have the following.

Corollary 2 *Let $\gamma = 1$ and consider a "deep PTA" that, in addition to eliminating bilateral tariffs, marginally increases bilateral recognition of IPRs, θ. Such deep provision enhances the welfare impact of the PTA (i.e., is a social strategic complement to bilateral tariff liberalization) if the existing level of bilateral IPRs is relatively low: $\theta < \alpha^O/\alpha$. Conversely, the deep provision reduces the welfare impact of the PTA (i.e., is a social strategic substitute to bilateral tariff liberalization) if the existing level of bilateral IPRs is relatively high: $\theta > \alpha^O/\alpha$.*

Hence, our model implies that "deeper" PTA provisions improve the impact of preferential tariff liberalization when IPRs are weak, but may not otherwise.

Another way of looking at the impact of deep provisions in PTAs is to consider how they affect the threshold $\bar{\omega}$. It is not difficult to see that $\bar{\omega}$ is concave in θ and reaches a maximum at an intermediate value of θ, $\hat{\theta} = \frac{2c - \sqrt{2c(2c-b^2)}}{\alpha b^2}$. Thus, deep integration amplifies the range of suppliers for which tariff preferences bring welfare gains whenever initial levels of investment protection are sufficiently low. Otherwise, deep integration shrinks the number of Y-chains for which the PTA increases welfare.

Analogously, we can see how the strength of IPRs affects the level of tariff preference consistent with the PTA being welfare-improving. See the Appendix for the proof.

Proposition 9 *When $\gamma = 1$, the highest level of the external tariff consistent with the PTA being welfare-improving, $2\hat{t}$, reaches a maximum at an interior level of IPRs, $\hat{\theta} = \frac{2c - \sqrt{2c(2c-b^2)}}{\alpha b^2}$.*

Thus, deep integration extends the level of the external tariff under which the PTA brings welfare gains whenever initial levels of investment protection are sufficiently low. Put differently, when either the fundamental hold-up problem is severe or IPRs are weak, deep integration is a social strategic *complement* to "shallow" integration, enhancing the efficacy of the tariff preference in promoting efficiency-enhancing investment. On the other hand, when investment protection is high, deep integration reduces the maximum level of the external tariff consistent with the PTA increasing welfare. Deep integration becomes then a social strategic *substitute* to shallow integration. In that case, there is a rationale for keeping the agreement restricted to its basic role of eliminating bilateral tariffs.

Observe that developing countries are typically associated with high tariffs (and high external tariffs under a PTA) and weak recognition of international IPRs (and therefore a low θ and a resulting low α'). This tends to generate conditions unfavorable to shallow integration (in the sense that t tends to be higher than $2\hat{t}$, since a low θ reduces $\hat{t}$). The introduction of deep provisions could therefore help to make "South-South" and "North-South" PTAs welfare-improving. Figure 8 illustrates that point.[28] When t is high and α' is low, $\Delta W(1) < 0$. If, however, the agreement also promotes a sufficiently large increase in α' (through an increase in θ), then $\Delta W(1) > 0$ becomes possible.

---

[28] The figure uses the same parametrization used in Figure 5, for $k = 2$. We note that it is well known that Pareto provides a good fit for the distribution of firm productivity in many contexts. This is the conclusion of, for example, the cross-industry analysis of Corcos et al. (2012) for the European Union. In particular, in their study the average parameter k across industries is estimated to very close to 2.
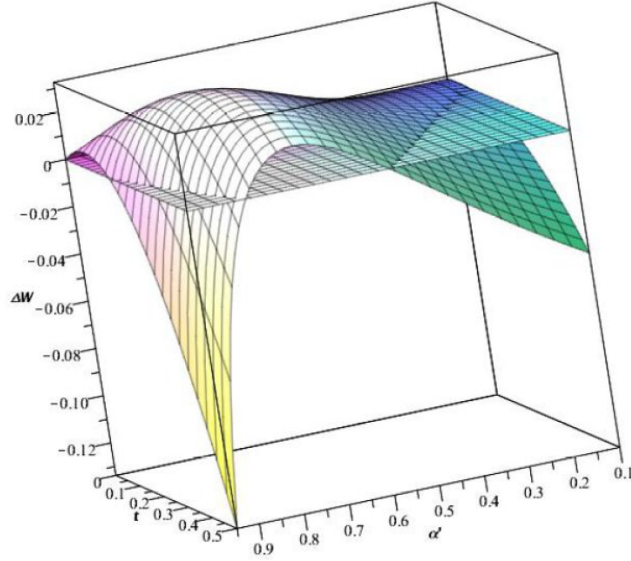
Fig. 8: $\Delta W(\gamma = 1)$ for Pareto ($k = 2$)

In contrast, developed economies are typically associated with low tariffs (and low external tariffs under a PTA) and strong IPRs regimes (and therefore a high θ and a resulting high α'). While this tends to provide generally favorable conditions for preferential liberalization (in the sense that t tends to be lower than $2\hat{t}$), our analysis suggests that "North-North" PTAs may be more effective if kept shallow. To see this in Figure 8, observe that, for combinations of very low t and very high α', $\Delta W(1) > 0$. However, if the agreement included deep provisions that induced a higher α', the welfare gain would not be as large.

At the cost of introducing some ambiguity in the results, one can readily extend the analysis to the general case where $\gamma \in [0,1]$. An important issue when doing that is to define whether the change in IPRs is indeed bilateral, only with respect to *Foreign*, or multilateral. Indeed, many deep provisions in recent PTAs do not have a preferential nature. Here we hint at what would be the additional effects of a deep PTA when the deep provision is not discriminatory.

Observe first that, when IPRs are nondiscriminatory, none of the matching cutoffs $\{\widetilde{\omega}_N, \widetilde{\omega}_F, \widetilde{\omega}_{ROW}\}$ depend upon α'. Then the analysis of how α' affects new suppliers is entirely analogous to the analysis of how it affects incumbent suppliers. The only important difference is that, because of the new suppliers' worse distribution of productivity, the level of supplier bargaining power that would maximize welfare for this group would be strictly below α^O.

46

On the other hand, the effect of α on $MD(\gamma)$ is entirely different: it can be shown that the welfare loss due to matching diversion is more severe, the higher is the supplier investment protection. This happens because the surplus generated by a Y-chain exhibits complementarity between productivity and supplier investment protection. As a result, the loss due to the reallocation of suppliers is especially large when suppliers have more bargaining power.

Hence, the effect of supplier investment protection on $NS(\gamma)$ also has two components: one has an inverse-U shape akin to the effect on $IS(\gamma)$, but shifted to the left; the other is negative and strictly decreasing. The net result is generally undefined because the density $f(\omega)$ could yield convex portions in $MD(\gamma)$. Barring very particular distributions, however, the α' that maximizes $NS(\gamma)$ will tend to be lower than α^O, but a similar analysis would carry through.

8   Positive Implications of a PTA

The main goal of our analysis is to investigate the welfare implications of PTAs under global sourcing. However, 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 whether a buyer is matched with a supplier in *Foreign* or in *ROW* prior to the PTA.

Specifically, we have that *buyers forming vertical chains in PTA member countries* prior to the agreement keep their original suppliers and source more from them. Thus, there is an *intensive margin positive effect* for incumbent suppliers in *Foreign*. Moreover, because of the higher investment levels, the productivity of those suppliers increases, so there is also a *productivity effect* for those matches.

Now, for *firms forming vertical chains in non-PTA countries* 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 switch to suppliers within the trading bloc, and their baseline productivity is lower than the productivity of their previous suppliers outside the bloc. Hence, there is also an *across-country extensive margin effect*, from outside to inside the trading bloc, for buyers originally matched with suppliers located outside the bloc that are not very productive.

Observe that the new suppliers inside the bloc did not invest and did not export before the

PTA. Therefore, the increase in investment is especially large for them and takes place only because they anticipate exporting. Interestingly, that testable prediction—i.e., a particularly large increase in investment for average-productivity producers that start to export *because* of preferential market access—is exactly what Lileeva and Trefler (2010) find in the context of preferential liberalization between Canada and the United States.

Recently, datasets that include the identity and characteristics of matched firms across countries are becoming increasingly available. If a PTA is implemented between two of the countries for which such data are available, one could investigate the validity of those relatively straightforward implications.

Sugita, Teshima and Seira (2018) provide an interesting analysis along those lines, but focusing 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.

9   Conclusion

Under global sourcing with incomplete contracts and endogenous buyer-supplier matching, a PTA affects the efficiency of the production process both through cost-reducing investment and through changes in the set of vertical chains. For that reason, a PTA can be welfare-enhancing even when there is no standard trade creation, as long as specialized suppliers are sufficiently productive and the tariff preference is not too high. The primary channel for positive welfare effects is through improved investments by suppliers originally located in PTA member countries. New vertical chains

could enhance welfare in circumstances where PTA countries have a large number of relatively productive suppliers that are idle without the PTA. However, rematching always lowers the average baseline productivity of suppliers and some new vertical chains always lower welfare under the PTA.

Deep provisions in PTAs enhance trade flows between members, but their welfare implications are subtle. For example, improved IPRs enhance investment protection, boosting incentives for relationship-specific investments. But that can improve or worsen the welfare impact of a PTA, depending upon whether changes in investment are already too strong under shallow integration. For that reason, shallow integration may be best for "North-North" agreements, whereas deep integration tends to be helpful for PTAs that involve developing economies where IPRs are lacking.

Our work is a small but we believe an important step toward understanding the implications of preferential liberalization in the context of global sourcing. In particular, our model offers a promising framework for future work. For example, one could extend the model to capture the effects of other deep provisions like improved product-quality standards. This could be modeled as an improved ability of a supplier to have the outside option to sell its output to firms other than her matched buyer. One could also adjust the model to capture the possibility that deep PTA provisions may select on productivity. If firms were required to pay fixed costs to take advantage of improved IPRs, say, then only higher-productivity firms would choose to do so. Hence, deep provisions could effectively achieve exclusion through facilitating choices that firms make. This has potential for framing empirical analyses of whether and how deep provisions select on productivity.

Our analysis also has implications for the design of PTAs. Studying further the optimality of preferential margins and of deep provisions is one natural way to proceed. Another is to consider criteria for selecting industries for exclusion from PTAs. Industry exclusion is a staple of PTAs. Although Article XXIV of the GATT requires that "substantially all trade" must be included in every preferential agreement, the vagueness of the requirement allows for very flexible interpretations. Furthermore, PTAs that do not include developed economies can be notified to the World Trade Organization under the "Enabling Clause," which imposes even weaker constraints, as Ornelas (2016) points out. As a result, in reality PTA exclusions vary from a few products to several entire sectors.[29] Surprisingly, there are very few theoretical analyses of sector exclusions in PTAs,

---

[29]For example, Deardorff and Sharma (2018) study 240 importer-exporter pairs within PTAs initiated prior to 2005, and find that the fraction of excluded products during 2009-11 is ten percent for the US and ranges between 3% and 44% for all countries in the sample.

the most notable exception being the political-economy analysis of Grossman and Helpman (1995). Here, we find that the high-productivity industries are the most valuable in an agreement if we considered only incumbent suppliers. However, once we take into account the influx of new suppliers, that conclusion is no longer warranted. Indeed, if it were feasible, a social planner would like to prevent, in every PTA and in every industry, the full market-driven reallocation of buyers across vertical chains.

Our framework could potentially 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 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. Alternatively, since the ROOs tend to increase the cost of production of North American firms by endogenously raising the cost of foreign inputs, one could model their tightening as a downward shift (possibly nonuniform) of the underlying distribution of productivity of suppliers within NAFTA. According to our analysis, that would lower the welfare impact due to all incumbent specialized suppliers. On the other hand, by worsening the distribution of productivity inside the bloc relative to the distribution of productivity outside the bloc, the tariff preferences *cum* tight ROOs may not yield any new suppliers, and instead generate "negative matching diversion"—with possible positive welfare effects.

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 (2017).

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. Critically, deep provisions in PTAs interact with preferential market access in a way that reinforces the latter's positive effect on trade flows but whose welfare implications are much more intricate than a simple look at trade flows would suggest. Thus, one could view our model as a step towards a framework that extends the Vinerian view to the "new regionalism" world.

Appendix

Efficient investment levels   Without an agreement, the efficient investment level solves

$$\max_i p_w q_N - C(q_N, i, \omega) - I(i). \tag{38}$$

The first-order necessary condition is

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

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

With a PTA, the efficient investment level also solves (38), after replacing q_N with q_P. The first-order necessary condition is analogous to the one above, but using (15) it simplifies to

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

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 $i = i^e$.

Explicit expressions for welfare Inserting equilibrium investments and levels of inputs, we have the following expressions for welfare:

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

Observe that the term in brackets in constant across trade regimes.

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.

No PTA Consider first the no PTA case. Let the suppliers pay the buyers a matching fee $M : \Omega \times \{F, ROW\} \times [0, b] \to \mathbb{R}$. Let the assignment of matches follow $\mu_N : [0, \beta] \to \Omega \times \{F, ROW\}$. Define the gross utility for a buyer of type b matched with a supplier of type ω in country y as $U_B(b, \omega, y)$. Define the gross utility for a supplier of type ω in country y matched with a buyer of type b as $U_S(\omega, y, b)$. Three sets of conditions must hold:

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

$$
\max_{\{\omega, y\}} U_B(b, \omega, y) + M(\omega, y, b).
$$

Given the fee, buyers maximize profits over a choice of supplier (productivity ω and location y).

2. For each supplier $(\omega, y) \in \Omega \times \{F, ROW\}$, each buyer match $b \in \mu_N^{-1}(\Omega, \{F, ROW\})$ solves

$$
i) \ \max_{\{b\}} U_S(\omega, y, b) - M(\omega, y, b).
$$

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

suppliers ($\beta < \gamma$), there is an additional requirement:

$$ii) \ \max_{\{b\}} U_S(\omega, y, b) - M(\omega, y, b) \leq 0 \ if \ \mu_N^{-1}(\Omega, \{F, ROW\}) \ \text{is empty.}$$

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:

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

Statement 1 implies that $\frac{dM(\omega, y, b)}{d\omega} = -\frac{dU_B(b, \omega, y)}{d\omega}$, and because U_B is a constant function of y, $M(\omega, y, b)$ is also a constant function of y. Because U_S is a constant function of b, statement 2(i) implies that $M(\omega, y, b)$ is a constant function of b. Statement 2(ii) implies the marginal supplier earns exactly zero profit, and that this supplier's ω is the same in both countries. Hence, we drop the b and y arguments from all functions and define $U_S^N(\omega) \equiv U_S(\omega, y, b)$, $U_B^N(\omega) \equiv U_S(b, \omega, y)$ and $M_N(\omega) \equiv M(\omega, y, b)$. This redefined notation is consistent with the notation in the main text. Denoting $\widetilde{\omega}_N$ as a marginal supplier, statement 2(ii) implies $U_S^N(\widetilde{\omega}_N) = M_N(\widetilde{\omega}_N)$. Profit maximization then guarantees that $\widetilde{\omega}_F = \widetilde{\omega}_{ROW} = \widetilde{\omega}_N$ for the no PTA case. Statement 3 and the monotonicity of $U_S^N(\omega)$ then imply that $F(\widetilde{\omega}_N) = \beta$.

Given that $\frac{dM_N(\omega)}{d\omega} = -\frac{dU_B(\omega)}{d\omega}$, it follows that $M_N(\omega)$ is the sum of $U_B^N(\omega)$ and a constant term. We construct the fee by specifying $M_N(\omega) = -U_B^N(\omega) + k_N$. Returning to a marginal supplier, we can then write $U_S^N(\widetilde{\omega}) = -U_B^N(\widetilde{\omega}) + k_N$. Hence, we can solve for k_N and substitute into $M_N(\omega)$ to find

$$M_N(\omega) = U_S^N(\widetilde{\omega}_N) - \left[U_B^N(\omega) - U_B^N(\widetilde{\omega}_N) \right].$$

PTA Now consider the PTA case, in which the level of discriminatory protection $\tau \in \{0, t\}$ for a B-S pair maps one-to-one with the supplier's location y: for $y = F$, we have $\tau = t$; for $y = ROW$, we have $\tau = 0$. Let the suppliers pay the buyers a matching fee $M : \Omega \times [0, b] \times \{0, t\} \rightarrow \mathbb{R}$. Let

the assignment of matches follow $\mu_P : [0, \beta] \to \Omega \times \{F, ROW\}$. Define the gross utility for a buyer of type b matched with a supplier of type ω, where the match enjoys discriminatory protection via tariff $\tau \in \{0, t\}$, as $U_B(b, \omega, \tau)$. Define the gross utility for a supplier of type ω matched with a buyer of type b, where the match enjoys discriminatory protection via tariff $\tau \in \{0, t\}$, as $U_S(\omega, \tau, b)$. Three sets of conditions must hold:

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

$$\max_{\{\omega, \tau\}} U_B(b, \omega, \tau) + M(\omega, b, \tau)$$

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 $b \in \mu_P^{-1}(\Omega, \{F, ROW\})$ solves

$$i)\ \max_{\{b\}} U_S(\omega, \tau, b) - M(\omega, b, \tau).$$

Given the fee, suppliers maximize profits over a choice of buyer. Because there is an excess of suppliers $(\beta < \gamma)$, there is an additional requirement:

$$ii)\ \max_{\{b\}} U_S(\omega, \tau, b) - M(\omega, b, \tau) \leq 0\ if\ \mu_P^{-1}(\Omega, \{F, ROW\})\ \text{is empty.}$$

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:

$$
\begin{aligned}
\int_{\mu_P([0,\beta] \times \{F, ROW\})} dF(\omega) &= \beta, \\
\int_{\mu_P([0,\beta] \times F)} dF(\omega) &\leq \gamma, \\
\int_{\mu_P([0,\beta] \times ROW)} dF(\omega) &\leq 1 - \gamma.
\end{aligned}
$$

Statement 1 implies that $\frac{dM(\omega, b, \tau)}{d\omega} = \frac{-dU_B(b, \omega, \tau)}{d\omega}$. Hence, $\frac{dM(\omega, b, 0)}{d\omega} = \frac{-dU_B(b, \omega, 0)}{d\omega}$ and $\frac{dM(\omega, b, t)}{d\omega} = \frac{-dU_B(b, \omega, t)}{d\omega}$. Because U_S is a constant function of b, statement 2(i) implies that $M(\omega, b, \tau)$ is a

constant function of b. Statement 2(ii) implies the marginal supplier earns exactly zero profit. Hence, we drop the b arguments from all functions and drop the tariff argument from functions when $\tau = 0$. Note that the gross payoffs for the buyer and supplier are, for the $\tau = 0$ case, the same as in the no-PTA case. In converting notation, we therefore write gross utilities under $\tau = 0$ as $U_S^N(\omega) \equiv U_S(\omega, 0, b)$ and $U_B^N(\omega) \equiv U_S(b, \omega, 0)$. For the $\tau = t$ case, we write gross utilities as functions of the tariff size, $U_S^P(\omega, t) \equiv U_S(\omega, t, b)$ and $U_B^P(\omega, t) \equiv U_S(b, \omega, t)$. For the fees, we write $M_{P,ROW}(\omega) \equiv M(\omega, b, 0)$ and $M_{P,F}(\omega, t) \equiv M(\omega, b, t)$.

Denoting $\widetilde{\omega}_{ROW}$ and $\widetilde{\omega}_F$ as marginal suppliers in ROW and F, respectively, we have $U_S^N(\widetilde{\omega}_{ROW}) = M_{P,ROW}(\widetilde{\omega}_{ROW})$ and $U_S^P(\widetilde{\omega}_F, t) = M_{P,F}(\widetilde{\omega}_F, t)$. We also know from statement 1 that $U_B^N(\widetilde{\omega}_{ROW}) + M_{P,ROW}(\widetilde{\omega}_{ROW}) = U_B^P(\widetilde{\omega}_F, t) + M_{P,F}(\widetilde{\omega}_F, t)$, because otherwise some buyers would not be maximizing profits. Hence, substituting, we can write $U_B^N(\widetilde{\omega}_{ROW}) + U_S^N(\widetilde{\omega}_{ROW}) = U_B^P(\widetilde{\omega}_F, t) + U_S^P(\widetilde{\omega}_F, t)$. It then follows immediately from the multiplicative separability of the total profit from a $B - S$ pair that $\widehat{\omega}_F = \widehat{\omega}_{ROW} + t$.[30] Statement 3 implies that $\gamma F(\widetilde{\omega}_{ROW} + t) + (1 - \gamma)F(\widetilde{\omega}_{ROW}) = \beta$.

We also have, because $\frac{dM_{P,ROW}(\omega)}{d\omega} = \frac{-dU_B^N(\omega)}{d\omega}$, that $M_{P,ROW}(\omega)$ may be written as the sum of $-U_B^P(\omega)$ and a term that does not vary with ω. Similarly, $M_{P,F}(\omega, t)$ may be written as the sum of $U_B^P(\omega, t)$ and a term that does not vary with ω. We construct the fees by specifying $M_P(\omega) = -U_B^P(\omega) + k_{P,ROW}$ and $M_P(\omega, t) = -U_B^P(\omega, t) + k_{P,F}$. Returning to the marginal suppliers, we can then write $U_S^N(\widetilde{\omega}_{ROW}) = -U_B^N(\widetilde{\omega}_{ROW}) + k_{P,ROW}$ and $U_S^P(\widetilde{\omega}_F, t) = -U_B^P(\widetilde{\omega}_F, t) + k_{P,F}$. Hence, we can solve for $k_{P,ROW}$ and $k_{P,F}$, then substitute to find

$$
\begin{aligned}
M_{P,ROW}(\omega) &= U_S^N(\widetilde{\omega}_{ROW}) - \left[U_B^N(\omega) - U_B^N(\widetilde{\omega}_{ROW}) \right], \\
M_{P,F}(\omega, t) &= U_S^P(\widetilde{\omega}_F, t) - \left[U_B^P(\omega, t) - U_B^P(\widetilde{\omega}_F, t) \right].
\end{aligned}
$$

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

Rewriting $NS(\gamma)$ using a change of variables   Start with the expression for the new supplier effect:

$$
NS(\gamma) \equiv \gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F(t)} \Psi_P(\omega, t) f(\omega) d\omega - (1 - \gamma) \int_{\widetilde{\omega}_{ROW}(t)}^{\widetilde{\omega}_N} \Psi_N(\omega) f(\omega) d\omega.
$$

---

[30] This is easily seen by plugging into (20) and (19) for both $\tau = 0$ and $\tau = t$, and adding the expressions together.

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

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

Then we note that

$$\gamma f(\widetilde{\omega}_F(x))d\omega = \frac{\phi(x;\gamma,F)d\omega}{d\widetilde{\omega}_F(x)} = \phi(x;\gamma,F)dx,$$

where the first equality follows from

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

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

$$\gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_F} \Psi_P(\omega,t)f(\omega)d\omega = \int_0^t \Psi_P(\widetilde{\omega}_F(x),t)\phi(x;\gamma,F)dx.$$

A similar manipulation of the second term in $NS(\gamma)$ yields

$$(1-\gamma) \int_{\widetilde{\omega}_{ROW}(t)}^{\widetilde{\omega}_N} \Psi_N(\omega)f(\omega)d\omega = \int_0^t \Psi_N(\widetilde{\omega}_{ROW}(x))\phi(x;\gamma,F)dx.$$

Hence,

$$NS(\gamma) = \int_0^t \left[\Psi_P(\widetilde{\omega}_F(x),t) - \Psi_N(\widetilde{\omega}_{ROW}(x))\right] \phi(x;\gamma,F)dx.$$

Proofs

Proof of Lemma 4. At $t = 0$, $\Delta\Psi(\omega,t) = 0$ by construction. We need to show, then, that a small increase in t, starting at $t = 0$, raises $\Delta\Psi(\omega,t)$. It is straightforward to see from (28) that $\frac{d\Delta\Psi_S(t=0)}{dt} = 0$. Now, we have that

$$
\begin{aligned}
\frac{d\Delta\Psi_R}{dt} &= \frac{2c-b^2}{2c}\left[(HUP_N - EXC_P)\frac{d\Delta i}{dt} - \Delta i\frac{dEXC_P}{dt}\right] \\
&= \frac{\alpha b\left(2c-b^2\right)}{2c\left(2c-\alpha b^2\right)}\left[HUP_N - EXC_P - \Delta i\right].
\end{aligned}
$$

Evaluated at $t = 0$, $\Delta i(t = 0) = 0$ and $EXC_P(t = 0) = i_N^* - i^e = -HUP_N$. Hence,

$$\frac{d\Delta\Psi_R}{dt}(t = 0) = \frac{\alpha b\left(2c - b^2\right)}{c\left(2c - \alpha b^2\right)}HUP_N > 0.$$

It follows that $\frac{d\Delta\Psi(\omega,t)}{dt}(t = 0) = \frac{d\Delta\Psi_R}{dt}(t = 0) > 0.$ ∎

Proof of Proposition 1. Equilibrium matching when $\gamma = 1$ requires

$$F_1(\widetilde{\omega}_{N1}) = \beta,$$

$$F_2(\widetilde{\omega}_{N2}) = \beta.$$

If $F_2(\omega)$ **FOSD** $F_1(\omega)$, the two distributions satisfy $F_1(\omega) \geq F_2(\omega)$. It follows that

$$\widetilde{\omega}_{N1} \leq \widetilde{\omega}_{N2}.$$

The changes in welfare from the PTA for the two distributions are

$$\Delta W_1(\gamma = 1; F_1) = \int_0^{\widetilde{\omega}_{N1}} \Delta\Psi(\omega,t)dF_1(\omega) \text{ and}$$

$$\Delta W_2(\gamma = 1; F_2) = \int_0^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega).$$

Hence,

$$\Delta\Delta W \equiv \Delta W_1(\gamma = 1; F_1) - \Delta W_2(\gamma = 1; F_2) = \int_0^{\widetilde{\omega}_{N1}} \Delta\Psi(\omega,t)dF_1(\omega) - \int_0^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega).$$

Integrating both terms by parts, we can write

$$
\begin{aligned}
\Delta\Delta W &= \Delta\Psi(\omega,t)F_1(\omega)\big|_0^{\widetilde{\omega}_{N1}} - \int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_1(\omega)d\omega - \left[\Delta\Psi(\omega,t)F_2(\omega)\big|_0^{\widetilde{\omega}_{N2}} - \int_0^{\widetilde{\omega}_{N2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_2(\omega)d\omega\right] \\
&= \Delta\Psi(\widetilde{\omega}_{N1},t)F_1(\widetilde{\omega}_{N1}) - \int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_1(\omega)d\omega - \left[\Delta\Psi(\widetilde{\omega}_{N2},t)F_2(\widetilde{\omega}_{N2}) - \int_0^{\widetilde{\omega}_{N2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_2(\omega)d\omega\right] \\
&= \beta\left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\widetilde{\omega}_{N2},t)\right] - \int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[F_1(\omega) - F_2(\omega)]d\omega + \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_2(\omega)d\omega \\
&= \left\{\beta\left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\widetilde{\omega}_{N2},t)\right] + \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \frac{d\Delta\Psi(\omega,t)}{d\omega}F_2(\omega)d\omega\right\} - \int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[F_1(\omega) - F_2(\omega)]d\omega.
\end{aligned}
$$

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

$$-\int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[F_1(\omega) - F_2(\omega)]d\omega > 0.$$

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

$$
\begin{aligned}
\{\cdot\} &= \beta\left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\widetilde{\omega}_{N2},t)\right] + \Delta\Psi(\omega,t)F_2(\omega)\big|_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} - \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega) \\
&= \beta\left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\widetilde{\omega}_{N2},t)\right] + \Delta\Psi(\widetilde{\omega}_{N2},t)F_2(\widetilde{\omega}_{N2}) - \Delta\Psi(\widetilde{\omega}_{N1},t)F_2(\widetilde{\omega}_{N1}) - \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega) \\
&= \beta\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\widetilde{\omega}_{N1},t)F_2(\widetilde{\omega}_{N1}) - \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega),
\end{aligned}
$$

where the final line comes from setting $F_2(\widetilde{\omega}_{N2}) = \beta$ and simplifying. We then have

$$
\begin{aligned}
\{\cdot\} &= \Delta\Psi(\widetilde{\omega}_{N1},t)\left[F_2(\widetilde{\omega}_{N2}) - F_2(\widetilde{\omega}_{N1})\right] - \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \Delta\Psi(\omega,t)dF_2(\omega) \\
&= \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\omega,t)\right]dF_2(\omega) > 0.
\end{aligned}
$$

Hence,

$$\Delta\Delta W = \int_{\widetilde{\omega}_{N1}}^{\widetilde{\omega}_{N2}} \left[\Delta\Psi(\widetilde{\omega}_{N1},t) - \Delta\Psi(\omega,t)\right]dF_2(\omega) - \int_0^{\widetilde{\omega}_{N1}} \frac{d\Delta\Psi(\omega,t)}{d\omega}[F_1(\omega) - F_2(\omega)]d\omega > 0,$$

concluding the proof. ∎

Proof of Proposition 2. By definition, the welfare impact of the PTA is zero when $t = 0$. When there is a small increase in t, $\Delta W(1)$ changes according to $\frac{\partial\Delta W(1)}{\partial t} = \int_0^{\widetilde{\omega}_N} \frac{\partial\Delta\Psi(\omega,t)}{\partial t}dF(\omega)$. We have that $\frac{\partial\Delta\Psi(\omega,t)}{\partial t} = \frac{2}{(2c-\alpha b^2)^2}\left\{-t\left[2c - 2\alpha b^2 + \alpha^2 b^2\right] + (p_w - \omega)\alpha(1-\alpha)b^2\right\}$. This expression is strictly positive when evaluated at $t = 0$. Therefore, for sufficiently small preference margins, $\Delta W(1) > 0$. Now notice that $\frac{\partial^2\Delta W(1)}{\partial t^2} = \int_0^{\widetilde{\omega}_N} \frac{\partial^2\Delta\Psi(\omega,t)}{\partial t^2}dF(\omega) = -\int_0^{\widetilde{\omega}_N} \frac{2[2c - 2\alpha b^2 + \alpha^2 b^2]}{(2c-\alpha b^2)^2}dF(\omega) < 0$. Therefore, $\Delta W(1)$ is maximized when $\frac{\partial\Delta W(1)}{\partial t} = 0$. Simple algebra shows that this happens when $t = \hat{t}$, as defined in (33). Finally, after some manipulation it follows that, when $t = 2\hat{t}$, $\Delta W(1) = 0$. Since $\frac{\partial^2\Delta W(t,1)}{\partial t^2} < 0$, $\Delta W(1) < 0$ when $t > 2\hat{t}$. ∎

Proof of Proposition 3. Since $\frac{d\Psi_N(\omega)}{d\omega} < 0$, we have that

$$\int_{\widetilde{\omega}_N}^{\widetilde{\omega}_{ROW}+t} \Psi_N(\omega)dF(\omega) < \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_{ROW}+t} \Psi_N(\widetilde{\omega}_N)dF(\omega)$$

and

$$\int_{\widetilde{\omega}_{ROW}}^{\widetilde{\omega}_N} \Psi_N(\omega)dF(\omega) > \int_{\widetilde{\omega}_{ROW}}^{\widetilde{\omega}_N} \Psi_N(\widetilde{\omega}_N)dF(\omega).$$

Now notice that

$$\gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_{ROW}+t} \Psi_N(\widetilde{\omega}_N)dF(\omega) - (1-\gamma) \int_{\widetilde{\omega}_{ROW}}^{\widetilde{\omega}_N} \Psi_N(\widetilde{\omega}_N)dF(\omega)$$

$$= \Psi_N(\widetilde{\omega}_N) \left[\gamma F(\widetilde{\omega}_{ROW}+t) + (1-\gamma)F(\widetilde{\omega}_{ROW}) - F(\widetilde{\omega}_N) \right]$$

$$= \Psi_N(\widetilde{\omega}_N) \left[\beta - \beta \right] = 0,$$

where in the last line we use the equilibrium conditions (14) and (22). Hence,

$$\gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_{ROW}+t} \Psi_N(\omega)dF(\omega) \quad < \quad \gamma \int_{\widetilde{\omega}_N}^{\widetilde{\omega}_{ROW}+t} \Psi_N(\widetilde{\omega}_N)dF(\omega)$$

$$= \quad (1-\gamma) \int_{\widetilde{\omega}_{ROW}}^{\widetilde{\omega}_N} \Psi_N(\widetilde{\omega}_N)dF(\omega) < (1-\gamma) \int_{\widetilde{\omega}_{ROW}}^{\widetilde{\omega}_N} \Psi_N(\omega)dF(\omega),$$

confirming that $MD < 0$. ∎

Proof of Proposition 4. Suppose $t > \frac{2\alpha(1-\alpha)b^2 \left[p_w - F^{-1}(\beta) \right]}{2c - 2\alpha b^2 + \alpha^2 b^2}$. Then

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

By Lemma 5, it follows that $NS(\gamma) = \int_0^t r(x,t)\phi(x;\gamma,F)dx < 0$. ∎

Proof of Proposition 5. We use $NS(\gamma) = \int_0^t r(x,t)\phi(x;\gamma,F)dx$. It is obvious that if $t = 0$, then $NS(\gamma) = 0$. Differentiating, we have

$$\frac{dNS(\gamma)}{dt} = r(t,t)\phi(t;\gamma,F) + \int_0^t \frac{dr(x,t)}{dt}\phi(x;\gamma,F)dx.$$

Because $r(0,0) = 0$, it is also obvious that $\frac{dNS(\gamma;t=0)}{dt} = 0$. Then, if $\frac{d^2NS(\gamma)}{dt^2} < 0$ for all t, then $NS(\gamma) < 0$ for all t as well. We now show that, under Condition 1, $\frac{d^2NS(\gamma)}{dt^2} < 0$ for all t. After

using the functional form for the r function to substitute, we have:

$$\frac{d^2 NS(\gamma)}{dt^2} = \left\{ \phi(t;\gamma,F) \left[\left(\frac{d}{dt} \frac{-2t\,(p_w - \widetilde{\omega}_{ROW}(t))}{(2c - \alpha b^2)} \right) + \frac{d\Psi'_P(t,\omega)}{dt} \right] \right\}$$
$$+ \int_0^t \frac{d^2 r(x,t)}{dt^2} \phi(\gamma,x;F)dx - \left[\frac{2t\,(p_w - \widetilde{\omega}_{ROW}(t))}{(2c - \alpha b^2)} \left(\frac{d\phi(t;\gamma,F)}{dt} \right) \right]. \qquad (39)$$

Start with the term in braces, expand the expression and substitute according to the functional form for $\frac{d\Psi'_P(t,\omega)}{dt}$:

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

Rearranging, we can write

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

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 (39). This is the aggregate of the second-order effects of the tariff for reallocations, each of which is negative. Hence, $\int_0^t \frac{d^2 r(x,t)}{dt^2} \phi(\gamma,x;F)dx < 0$. Finally, consider the second term on the second line of (39). Because $\frac{d\phi(t;\gamma,F)}{dt} \geq 0$ under Condition 1, the entire term is non-positive. This shows that $\frac{d^2 NS(\gamma)}{dt^2} < 0$. $\blacksquare$

Proof of Proposition 6. Define $\underline{t}$ to be the lowest value of t such that $\Delta W(\gamma) = 0$. Differentiating, we have that $\frac{d\Delta W(\gamma)}{dt} = \frac{dIS(\gamma)}{dt} + \frac{dNS(\gamma)}{dt}$. In the limit, $\lim_{t \to 0} \frac{dIS(\gamma)}{dt} > \lim_{t \to 0} \frac{dNS(\gamma)}{dt} = 0$. Hence, $\lim_{t \to 0} \frac{d\Delta W(\gamma)}{dt} > 0$ and $\underline{t} > 0$.

From Proposition 2, $IS(\gamma) < 0$ for any $t > 2\hat{t}$ and $IS(\gamma)$ is decreasing in t for any $t > \hat{t}$. From Proposition 4, $NS(\gamma) < 0$ for any $t > t^{NS}$. It is straightforward to show that $\hat{t} < t^{NS} < 2\hat{t}$. Hence, if $t \geq 2\hat{t}$, then $\Delta W(\gamma) = IS(\gamma) + NS(\gamma) < 0$. By continuity of $\Delta W(\gamma)$, it follows that $\Delta W(\gamma) < 0$ for some $t < 2\hat{t}$ as well. Define $\bar{t}$ to be the highest t such that $\Delta W(\gamma) = 0$. Thus, we have shown

that $\bar{t} \in [\underline{t}, 2\hat{t})$.

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

Proof of Proposition 7. Let $\gamma < 1$. We will show that ΔW is strictly decreasing in t for any $t \geq \hat{t}$. We can write $\Delta W = \gamma \int_0^{\widetilde{\omega}_N} \Delta\Psi(\omega, t)dF(\omega) + NS(\gamma)$. From Proposition 2 we know that $\int_0^{\widetilde{\omega}_N} \Delta\Psi(\omega, t)dF(\omega)$ is maximized at $t = \hat{t}$ and has an inverted-U shape with respect to t. Hence, the derivative of γ times this term with respect to t is zero at $t = \hat{t}$ and is negative for $t > \hat{t}$

Let $t \geq \hat{t}$. Recall that

$$rw(0,t) = \frac{t}{(2c - \alpha b^2)^2} \left[2b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_N) - t(2c + \alpha^2 b^2 - 2\alpha b^2) \right].$$

Differentiating, we have

$$
\begin{aligned}
\frac{rw(0,t)}{dt} &= \frac{1}{(2c - \alpha b^2)^2} \left\{ \left[2b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_N) - t(2c + \alpha^2 b^2 - 2\alpha b^2) \right] - t(2c + \alpha^2 b^2 - 2\alpha b^2) \right\} \\
&= \frac{1}{(2c - \alpha b^2)^2} \left[2b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_N) - 2t(2c + \alpha^2 b^2 - 2\alpha b^2) \right],
\end{aligned}
$$

which is negative if
$$t > \frac{b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_N)}{(2c + \alpha^2 b^2 - 2\alpha b^2)}.$$

Note that
$$\hat{t} = \frac{\alpha(1-\alpha)b^2 \left[p_w - \mathbf{E}\left(\omega; \omega \leq \widetilde{\omega}_N\right) \right]}{2c - 2\alpha b^2 + \alpha^2 b^2} \geq \frac{b^2\alpha(1-\alpha)(p_w - \widetilde{\omega}_N)}{(2c + \alpha^2 b^2 - 2\alpha b^2)}.$$

Hence, if $t \geq \hat{t}$, then $\frac{drw(0,t)}{dt} < 0$. Now, we can also show that $\frac{drw(x,t)}{dt}$ is decreasing in x:

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

This implies that
$$\int_0^t \frac{drw(x,t)}{dt}\phi(x; \gamma, F)dx < 0.$$

Because $\phi(t; \gamma, F)rw(t,t) < 0$ for any t, we have

$$NS'(t) = \phi(t; \gamma, F)rw(t,t) + \int_0^t \frac{drw(x,t)}{dt}\phi(x; \gamma, F)dx < 0.$$

This show that $NS(t)$ is decreasing in t for any $t \geq \hat{t}$. Hence, $\hat{t}$ does not maximize ΔW. $\blacksquare$

Proof of Proposition 8.   It follows immediately from (33), after replacing α by α', that $\lim_{\alpha' \to 0} \hat{t} = \lim_{\alpha' \to 1} \hat{t} = 0$. Therefore, since a PTA is defined by $t > 0$, $\lim_{\alpha' \to 0} \Delta W(1) < 0$ and $\lim_{\alpha' \to 1} \Delta W(1) < 0$. Simple algebra shows that $\lim_{\alpha' \to 0} \frac{\partial \Delta \Psi(\omega, t)}{\partial \alpha'} > 0$ and $\lim_{\alpha' \to 1} \frac{\partial \Delta \Psi(\omega, t)}{\partial \alpha'} < 0$; hence, $\lim_{\alpha' \to 0} \frac{\partial \Delta W(1)}{\partial \alpha'} > 0$ and $\lim_{\alpha' \to 1} \frac{\partial \Delta W(1)}{\partial \alpha'} < 0$. Now, setting $\frac{\partial \Delta W(1)}{\partial \alpha'} = 0$ and manipulating, we obtain a single solution for α', given by expression (37). Since $\Delta W(1)$ is increasing in α' when α' is close to one but decreasing in α' when α' is close to zero, α^O must define a maximum. $\blacksquare$

Proof of Proposition 9.   After replacing α by $\alpha' = \alpha\theta$, differentiate (33) with respect to α' and reorganize to obtain

$$\frac{\partial \hat{t}}{\partial \alpha'} = -\frac{2b^2 \left(2c - 4\alpha'c + (\alpha')^2 b^2 \right) [p_w - \mathbf{E}(\omega; \omega \leq \widetilde{\omega}_N)]}{\left[2c - 2\alpha'b^2 + (\alpha')^2 b^2 \right]^2}.$$

Solving this expression for θ yields $\hat{\theta} = \frac{2c - \sqrt{2c(2c - b^2)}}{\alpha b^2}$ as the unique stationary point of the function $\hat{t}(\theta)$. Since we know that $\hat{t} > 0$ except at the extreme values of α', when it is zero, $\hat{\theta}$ must constitute a maximum of $\hat{t}(\theta)$. $\blacksquare$

References

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

Antràs, Pol and Alonso de Gortari (2017). "On the Geography of Global Value Chains." Mimeo.

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.

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.

Boehm, Johannes and Ezra Oberfield (2018). "Misallocation in the Market for Inputs: Enforcement and the Organization of Production." Mimeo.

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.

Corcos, Gregory, Massimo Del Gatto, Giordano Mion and Gianmarco Ottaviano (2012). "Productivity and Firm Selection: Quantifying the new Gains from Trade." *Economic Journal* 122(561), 754-798.

Crivelli, Pramila (2016). "Regionalism and falling external protection in high and low tariff members." *Journal of International Economics* 102, 70-84.

Deardorff, Alan and Rishi Sharma (2018). "Exempted Sectors in Free Trade Agreements." Mimeo.

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.

Heise, Sebastian, Justin Pierce, Georg Schaur and Peter K. Schott (2015). "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.

Mattoo, Aaditya, Alen Mulabdic and Michele Ruta (2017). "Trade Creation and Trade Diversion in Deep Agreements." Working paper.

Ornelas, Emanuel (2016). "Special and Differential Treatment for Developing Countries." In: K. Bagwell and R. Staiger Eds.), *Handbook of Commercial Policy*, Elsevier/North Holland, Amsterdam.

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.

Osnago, Alberto, Nadia Rocha and Michele Ruta (2018). "Deep Trade Agreements and Vertical FDI: The Devil is in the Details." *Canadian Journal of Economics*, forthcoming.

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 2017, *Measuring and Analyzing the Impact of GVCs on Economic Development*, pp. 175-185. World Bank, Washington DC.

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

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

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

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