Preferential Trade Agreements and Global Sourcing*

Grant Bickwit       Emanuel Ornelas       John L. Turner

May 11, 2017

Abstract

We study how a preferential trade agreement (PTA) affects international sourcing decisions under incomplete contracting. The model features dual sourcing and search costs. Contract incompleteness implies underinvestment by the parties in a transaction. From a static standpoint, we show that this inefficiency is mitigated by a PTA, because the agreement allows the parties to internalize a larger fraction of their bargaining surplus. On the other hand, the PTA also yields sourcing diversion. A more efficient supplier tilts the tradeoff toward the beneficial investment effect; a higher external tariff tips it toward sourcing diversion. In a dynamic version of the model with search and matching, a PTA also has the benefit of inducing more search—which otherwise would be inefficiently low—prompting firms with mediocre matches outside the PTA to re-match with more productive suppliers inside the bloc. Hence, the dynamic consequences of a PTA, welfare-enhancing by themselves, also foster the circumstances under which the agreement tends to be welfare-improving from a static viewpoint. On the whole, we offer a new framework to study the benefits and costs from preferential liberalization in the context of global sourcing.

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

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

[Preliminary and incomplete]

*Bickwit: University of Georgia; Ornelas: Sao Paulo School of Economics-FGV, CEPR, CESifo and CEP-LSE; eaornelas@gmail.com; Turner: Department of Economics, University of Georgia; jlturner@uga.edu. We thank ...
1 Introduction

The past few decades have seen an explosion in the number of Preferential Trade Agreements (PTAs). The World Trade Organization (WTO) reports over 400 reciprocal PTAs in force in 2016, including all of its members, compared with just a few dozen in the early 1990s. There are also numerous schemes of unilateral trade preferences.\(^1\) A parallel trend has been the growth of trade in parts and components. As Johnson and Noguera (2016) document, the ratio of trade in value added to gross exports (which they call ‘the VAX ratio’) has declined steadily in the last 40 years. Increasingly, researchers are arguing that those two trends are related to each other (see for example Baldwin, 2011; World Trade Organization, 2011). Indeed, Johnson and Noguera (2016) find a strong link between reductions in the VAX ratio and the formation of PTAs. We still lack, however, a basic framework that allows us to assess the desirability of PTAs in that context. This is what we aim to provide in this paper.

Conventional welfare analyses of preferential liberalization typically point to two opposing effects of preferential tariffs, trade creation and trade diversion (Viner, 1950). 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 on market-clearing for price formation and neglect the nuances of real-world trade in intermediate goods. This is why some authors, like Baldwin (2011, 2016), have argued that 21\(^{st}\) century regionalism is no longer about preferential market access and the resulting trade creation/diversion, but mostly about the disciplines that underpin production fragmentation. Antràs and Staiger (2012) make a related point when studying the economics of (nondiscriminatory) trade agreements.

There are two key distinguishing features of trade in intermediates. First, it often involves customized components that commit a buyer and a seller to each other. It is well known that such bilateral monopoly can lead to underinvestment in component-specific technology due to ‘hold-up problems’ (e.g., Grossman and Hart, 1986). For example, by renegotiating terms of trade a buyer of customized components can hold up the seller and force a new bargain where she captures some

\(^1\)See https://www.wto.org/english/tratop_e/region_e/region_e.htm and for detailed statistics on preferential integration.
of the surplus created by sunk investments made by the seller. When contracts are incomplete, the seller anticipates the outcome of the renegotiation and underinvests. Second, finding an appropriate supplier involves costly search.

In this paper, we show that PTAs can enhance welfare both by (imperfectly) substituting for weak contract enforcement and by inducing efficiency-enhancing search effort. To make our message as stark as possible, we develop a simple model that shuts down all Vinerian trade creation channels but allows for trade diversion. We design the model to put aside classic trade creation not because we deem it unimportant; 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. Instead, we want to shed light on potentially important forces that have so far been ignored in the academic literature and in policy circles alike. Still, all the action is driven by preferential market access.

In our setting, in comparison with the case without an agreement, PTA trading partners share a higher surplus on every unit traded, relative to what they could obtain by dealing with alternative producers in nonmember countries. This propels firms to trade more, which in turn induces them to increase their relationship-specific investments. Since the investment yields fruits (i.e., greater value) to every unit traded, this relationship-strengthening effect is stronger, the more units the firms initially trade. Since without the PTA there is underinvestment due to a hold-up problem, the PTA-induced investment will tend to improve efficiency. This beneficial effect may overcome the negative effect of the tariff discrimination—essentially, trade diversion in the sourcing of components. This sourcing diversion is independent of by the number of units that the firms initially trade with each other. Therefore, the positive relationship-strengthening effect is more likely to dominate the negative sourcing-diversion effect when firms initially trade high volumes. This happens when the firms trading customized components are relatively efficient. Hence, we introduce a new element into Viner’s classic tradeoff and show that a PTA is more likely to enhance welfare when it is applied to more efficient industries, which trade large volumes of specialized inputs even without the PTA, in member countries where contract enforcement is fundamentally weak.2

2This 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.
Now, as the external tariff rises, the initially positive relationship-strengthening effect grows but eventually falls, whereas the negative sourcing-diversion effect increases monotonically in the tariff. Hence, if external tariffs are sufficiently high, PTAs are likely to lower welfare even in industries with highly productive firms. Thus, as in the classical case, with sufficiently high external tariffs trade diversion dominates. Yet recall that here the comparison is not with classic trade creation, but with the investment effect. If the discrimination is too high, it yields “too much” investment, possibly more than offsetting the benefit of alleviating the original hold-up problem.

Furthermore, specialized suppliers are not freely available to interested buyers. Instead, they need to be sought after through a costly search process. Introducing search and matching to the analysis allows us to study how PTAs affect the location of customized sourcing, while making productivity endogenous. A first finding is that without a PTA, there is too little search. Intuitively, this occurs for the same reason investments are too low: those incurring the search cost do not absorb the full benefits from the search process.

A PTA mitigates that inefficiency by inducing a higher search intensity. This yields matches with higher-productivity suppliers. But recall that the PTA tends to enhance welfare precisely when suppliers have high productivity. The PTA also induces firms initially sourcing elsewhere to seek new suppliers in the PTA members. The upshot is that the distribution of matches improves, both with suppliers in the PTA partner and with suppliers outside the PTA. This helps to rationalize the finding of Johnson and Noguera (2016) that PTAs help to explain the observed pattern of international sourcing. It also indicates a new channel through which PTAs promote efficiency.

Our paper illustrates how taking Viner’s (1950) distinction further can change the normative implications of PTAs, sometimes entirely reversing Viner’s original idea: we show that even purely trade-diverting PTAs can be helpful, when one considers how they can mitigate hold-up problems and search inefficiencies created by incomplete contracts.³ 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 investment and search incentives.

In that sense, our paper adds to the literature that seeks to link trade liberalization to investment

---

³In a way, this is the flip coin of the message from Ornelas (2005b), who shows that even purely trade-creating PTAs can be harmful, if one considers their implications for multilateral liberalization.
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 (Argentinean firms facing lower tariffs in Brazil under Mercosur in one case, Canadian firms facing lower tariffs in the U.S. under CUSTA in the other), although their models pay no heed to the preferential nature of the liberalization. In contrast, we highlight on the discriminatory aspect of tariff changes. Furthermore, we are interested in how they affect investment and matching related to international sourcing decisions, not a special concern in the analyses of Bustos (2011) and Lileeva and Trefler (2010).

Our analysis also provides a basic criterion 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 WTO under the "Enabling Clause," which imposes even weaker constraints. As a result, in reality PTA exclusions vary from a few products to several entire sectors. Surprisingly, there are very few theoretical analyses of sector exclusions in PTAs. The most notable exception is Grossman and Helpman (1995). According to their political-economy analysis, the welfare-enhancing sectors where trade creation would dominate are the ones that tend to be excluded, because those sectors do not receive support from organized lobbies. This tends to make actual (that is, politically viable) PTAs welfare-decreasing. Here, we find that the low-productivity industries are the ones that should be excluded, because in those sectors the mitigation of hold-up problems is less valuable.

Our analysis complements research using detailed models of intermediate input trade and bargaining in international trade (among others, Antràs and Helpman, 2004, 2008; Antràs and Chor, 2013; Antràs and Staiger, 2012; Ornelas and Turner, 2008, 2012). In terms of structure, we build on Ornelas and Turner (2008, 2012), but the current model departs from them in at least three

---

4 Ornelas (2005a) finds that this general conclusion may be reversed if the endogeneity of external tariffs is taken into account, but he does not consider exclusions.
fundamental ways. First, the focus is very different: those papers do not study preferential liberalization. Second, heterogeneity in productivity is a central feature of the current analysis, whereas Ornelas and Turner (2008, 2012) consider a representative firm. Third, search and matching are essential ingredients of the current analysis, but are absent in those papers.

The paper is also closely related to Antràs and Staiger (2012). Antràs and Staiger study optimal trade agreements in the presence of hold-up problems and prices negotiated through bargaining. We do not study optimal trade agreements, instead focusing on the impact of the (exogenous) introduction of a PTA. Nevertheless, the environments in the two papers are sufficiently similar to deserve a more detailed comparison. We offer that comparison in section 3.1. For now, it suffices to say that, while Antràs and Staiger (2012) focus on allocative efficiency, our concern is with dynamic efficiency. As we argue in section 3.1, this distinction is important for several reasons.

The paper is organized as follows. We setup 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 in place and describe the impact of the agreement on firms’ choices. We then study the welfare impact of the PTA (section 5). We introduce search costs and study how a PTA affects the location and the efficiency of sourcing in section 6. We conclude in section 7.

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 $x$ that enters consumers’ utility function linearly. Thus, if consumers purchase any amount of $x$, any extra income will be directed to the consumption of differentiated goods. We assume relative prices are such that consumers always purchase some good $x$. Furthermore, we assume that production of one unit of $x$ requires one unit of labor, that the market for good $x$ is perfectly competitive, and that $x$ 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 an intermediate input under conditions of decreasing returns to scale. Production is carried out by a Buyer ($B$) firm located in the Home country. Those firms act as
aggregators, transforming intermediate inputs into marketable goods. Final good producers obtain revenue \( V(Q) \) from purchasing a total of \( Q \) intermediate inputs, where \( V' > 0 \) and \( V'' < 0 \).

There is another country, \textit{Foreign}, as well as the rest of the world (\textit{ROW}). When sourcing, each buyer may purchase generic inputs available in the world market, \( g \), and/or customized inputs \( q \) from a specialized supplier (\( S \)), located in either \textit{Foreign} or \textit{ROW}. Generic inputs are produced by a competitive fringe and require \( p_w \) units of labor. Thus, their price in the world market is \( p_w \). \textit{Home}’s buyers face a per-unit tariff \( t \) on all imported intermediate goods, so a generic input costs \( p_w + t \) for them.\(^5\) Generally, a buyer will value generic and customized inputs differently. However, without loss of generality we can normalize units so that one unit of generic input and one of customized input have the same revenue-generating value for a buyer.\(^6\) 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 \( Q \).

Now, to acquire customized inputs, a buyer must first match with a supplier. To find a supplier with whom to match, a buyer must incur costly search in either \textit{Foreign} or \textit{ROW}. Upon finding a satisfactory match, \( B \) 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. As we solve the game by backward induction, initially we carry out the analysis for a given match. In section 6 we then return to the first stage of the game, the search and match analysis.

After a matching occurs and \( B \) and \( S \) specialize toward each other, \( S \) makes a non-contractible relationship-specific investment and pays for it. The benefit of \( S \)’s investment is that it lowers her marginal cost prior to trade with \( B \). Observe, however, that the exact nature of the investment benefit for the relationship is largely immaterial. For example, we could consider instead that \( S \)’s ex-ante investment affects not her cost function, but \( B \)’s valuation for specialized inputs, relative to generic ones (we make this claim precise momentarily, right after we define \( S \)’s technology). Similarly, nothing essential would change if the buyer, rather than the supplier, were the responsible

\(^5\)The analysis would remain virtually unchanged if we allowed for domestic production of generic inputs, provided that \textit{Home} remained an importer of generic inputs. Essentially, that would simply amount to a reinterpretation of \( B \)’s demand for inputs, \( V'(Q) \), as his demand for foreign inputs.

\(^6\)Specifically, we can add a multiplicative ‘compatibility cost’ to the use of generic inputs. Call such costs \( \xi \). This would increase the effective cost of generics for their buyers to \( \xi p_w + t \). But we can then simply redefine units by dividing the units of generic inputs by \( \xi \) and adjusting the tariff accordingly.
for the investment. All that matters is that there is an ex-ante—that is, prior to $B$-$S$’s choice of how many units of specialized inputs to trade—investment by either $S$ or $B$ (or by both) that enhances the value of trade within the relationship.

Once investment is sunk, the firms decide on how much to trade and at what price. The parties cannot use contracts to affect their trading decisions either. 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. As a result, the parties need to bargain over price and quantity of specialized inputs. If bargaining breaks down, $S$ produces nothing and $B$ has no choice but to purchase only generic inputs. If bargaining is successful, then $B$ purchases generic inputs from $ROW$ but also specialized inputs from $S$. Finally, $B$ transforms all inputs into the final good and payoffs are realized.

Each particular supplier is identified by \( \omega \), a heterogeneity parameter that indexes (the inverse of) her productivity. The specialized inputs are not traded on an open market, and therefore have no value outside the $B$-$S$ relationship. Conditional on investment $i$, the supplier’s cost function is

\[
C(q, i, \omega) = (\omega - bi)q + \frac{c}{2}q^2,
\]

where $q$ denotes her input production. Note that parameter $\omega$ shifts the firm’s marginal cost at rate 1. The lower is $\omega$, the more efficient the firm is. In turn, parameter $c$ determines the slope of the supplier’s marginal cost. Finally, parameter $b$ denotes the effectiveness of investment in reducing her production costs.

The investment is observed by both $B$-$S$, but is not verifiable in a court of law. Its cost is

\[
I(i) = i^2.
\]

Investment is bounded by $i \in [0, i^{max}]$. We assume that $2c > b^2$.\(^7\)

Those specific functional forms are not essential for our results. We impose them, to a large extent, to generate clear-cut analytical solutions. We would be able to generate similar results with\(^7\)

---

\(^7\)This ensures that the effect of investment on marginal cost is not too large relative to the elasticity of the cost function. If $b$ were too large, the supplier would want to make $i \to \infty$.\]
a more general specification for functions $C(.)$ and $I(.)$, given some restrictions on their second and third derivatives. Observe, nevertheless, that it is very common in such analyses to assume a constant marginal cost, and then to equate marginal cost with (the inverse of) productivity. Instead, we allow for the intercept of the marginal cost curve to depend on a productivity parameter and on ex-ante investment, in addition to allowing for a varying slope.

As indicated above, an alternative way to model investment is to consider that it enhances the value of customized inputs used by $B$. Specifically, considering that investment lowers the marginal cost of production, as we do, is isomorphic to a case where one unit of customized input is as valuable for $B$ as $1 + bi$ units of generic inputs. In that case, $S$’s cost structure does not change with investment, but her investment increases the efficiency of customized inputs used by $B$. Despite the different interpretation, the analysis would remain essentially unchanged under that alternative approach.

We focus on the case where $B$ engages in dual sourcing. To guarantee that $S$ produces some inputs, we need $C_q(0, 0, \omega) < p_w$ for all $\omega$. This is equivalent to assuming that $\omega < p_w$. Now define $Q^*$ as the level of total inputs sourced. This satisfies $V'(Q^*) = p_w + t$. To ensure production of the final good, the initial level of marginal revenue for $B$ needs to be sufficiently high: $V'(0) > p_w + t$. Now, to ensure that $S$ does not produce all inputs, we assume $C_q(Q^*, \max, 0) > p_w$, 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.

It is important to emphasize that dual sourcing is not essential for our results. In addition to being realistic, the main role of the dual sourcing specification is pedagogical, as will become clear in the analysis. In fact, most of the results can be extended to the case where there is single sourcing, but where the buyer can choose between purchasing customized or generic inputs. 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.

### 3 No Trade Agreement

Initially, there is no trade agreement and all inputs imported into $Home$ are subject to the tariff regardless of their origin. Conditional on investment $i$ and on the tariff, a pair $B$-$S$ trades the
ex-post privately efficient number of specialized inputs, and \( B \) purchases the ex-post efficient level of generic inputs. Efficient sourcing of customized and generic inputs, respectively \( q_N \) and \( g_N \), satisfies the following two conditions:

\[
Q^* = q_N + g_N, \tag{1}
\]

\[
C_q(q_N, i, \omega) = p_w. \tag{2}
\]

The latter condition implies

\[
q_N = \frac{p_w - \omega + bi}{c}. \tag{3}
\]

After \( S \) chooses her investment, \( B \) and \( S \) determine the price of the specialized intermediate inputs, \( p_s \), by Generalized Nash bargaining over the surplus due to trading customized instead of generic inputs. Specifically, let the supplier have bargaining power \( \alpha \in (0, 1) \). Under Generalized Nash bargaining, the two firms maximize the following expression with respect to \( p_s \), the price of the inputs they bargain over:

\[
(U^T_B - U^N_B)^{(1-\alpha)}(U^T_S - U^N_S)^\alpha,
\]

where \( U^J_k \) 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 (\( N \)). Those values are laid out as follows: \( U^T_B = V(Q^*) - (p_w + t)q_N - (p_s + t)q_N; U^0_B = V(Q^*) - (p_w + t)Q^*; U^T_S = p_sq_N - C(q_N, i, \omega); U^0_S = 0 \).

Defining \( \Omega \equiv (U^T_B - U^0_B) + (U^T_S - U^0_S) \) 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 the absence of a preferential trade agreement,

\[
\Omega_N = p_wq_N - C(q_N, i_N, \omega). \tag{4}
\]

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,

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

Note that as either the seller’s bargaining power ($\alpha$) or the effectiveness of investment ($b$) rises, investment increases. Conversely, a steeper marginal cost curve (a higher $c$) brings investment incentives down.

Substituting (5) back in (3) and manipulating, we find

$$q_N^* = \left( \frac{2}{\alpha b} \right) \left( \frac{\alpha b}{2c - \alpha b^2} \right) (p_w - \omega) \quad = \left( \frac{2}{\alpha b} \right) i_N^*. \quad (6)$$

Hence, the equilibrium investment and output are proportional. Thus, more productive firms produce more output for a given investment, and they also invest more. This reinforces their original advantages. When $\alpha$ is very small, the investment is very low, and drops to zero when $\alpha = 0$, when $S$ does not appropriate any of the benefits of her investment. 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 efficient sourcing, worldwide social welfare due to this bilateral transaction can be defined as

$$\Psi = V(Q^*) - p_w Q^* + p_w q_N - C(q_N, i, \omega) - I(i).$$

The efficient level of investment ($i^e$) maximizes $\Psi$. Observe that, given dual sourcing, the first two terms of $\Psi$ are unaffected by the level of investment. Thus, using (2), it follows that efficiency requires

$$I'(i^e) = -C_i(\cdot). \quad (6)$$

This yields

$$i^e = \left( \frac{b}{2c - b^2} \right) (p_w - \omega).$$
Observe that, as $b$ approaches $\sqrt{2c}$, the level of the efficient investment blows up.\(^8\) Comparing $i^*_N$ with $i^e$, it is immediate that $\alpha < 1$ yields $i^*_N < i^e$. Moreover, the difference $(i^e - i^*_N)$ increases if $S$ were more productive (i.e., if $\omega$ were lower). Thus, it is precisely the relationships with the best suppliers that are more negatively affected by contract incompleteness.

3.1 Comparison with the Antràs and Staiger’s (2012) framework

...  

**Observation:** unlike in AS, here price of final good has no effect on bargaining or investment decisions

**Reason:** dual sourcing.

This helps us to isolate effects due to discriminatory nature of PTA.

...

**Observation:** here, unlike in AS, firms bargain over their terms of exchange, *but* the bargaining is disciplined by market clearing conditions of alternative products.

...

4 A Preferential Trade Agreement

Under a PTA, the tariff for goods traded between *Home* and *Foreign* is eliminated. Imports from *ROW* still face the tariff, assumed unchanged. Hence, the specialized supplier’s price drops to $p_s$, as opposed to $p_s + t$ without the agreement.\(^9\) For expositional simplicity, we assume that *Foreign* does not produce generic inputs.\(^{10}\) Accordingly, generic inputs $g$ still cost $p_w + t$ for *Home*’s buyers.

The total volume of inputs purchased by $B$ remains unchanged at $Q^*$, as pinned down by $V'(Q^*) = p_w + t$, but now the composition of the sourcing decision changes to reflect the new relative prices. This is summarized by the condition

$$C_q(q_P, i^P, \omega) = p_w + t,$$

\(^8\)In this case, $i^{\text{max}}$ would obtain as a corner solution.

\(^9\)Naturally, the negotiating price $p_s$ also changes endogenously with the PTA.

\(^{10}\)Alternatively, we could assume that *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. This would leave all of our main results essentially unchanged.
which in equilibrium yields

\[ q_P = \frac{p_w + t - \omega + bi}{c}. \]  

(8)

Only one of the potential \( U_k^j \) terms, \( U_B^T \), structurally changes, becoming

\[ U_B^T = V(Q^*) - (p_w + t)q_P - p_s q_P. \]

The bargaining surplus under a trade agreement, \( \Omega_P \), 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 \( \Omega_P \) as they do without a trade agreement. Accordingly, the investment decision becomes the solution of

\[ \max_i \alpha \Omega_P - I(i_P). \]

The optimal investment under the PTA thus satisfies

\[ i_P^* = \left( \frac{ab}{2c - \alpha b^2} \right) (p_w + t - \omega). \]  

(9)

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

The resulting equilibrium level of customized inputs,

\[ q_P^* = \left( \frac{2}{\alpha b} \right) i_P^*, \]  

remains proportional to investment. Hence, the increases in both investment (\( \Delta i \)) and in customized inputs (\( \Delta q \)) are proportional to the size of the external tariff:

\[ \Delta i = i_P^* - i_N^* = \left( \frac{ab}{2c - \alpha b^2} \right) t, \]

\[ \Delta q = q_P^* - q_N^* = \left( \frac{2}{2c - \alpha b^2} \right) t. \]
Both of them increase with the seller’s bargaining power $\alpha$ and with the responsiveness of marginal cost to investment $b$, and fall with the slope of the marginal cost curve $c$.

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 supplier had no bargaining power ($\alpha = 0$), and therefore did not invest, $\Delta q(\alpha = 0) = \frac{t}{c} > 0$.

Under the PTA, $S$’s investment enhances the bargaining surplus that $B$ and $S$ share by more than it does without a trade agreement. Since $\alpha > 0$, $S$ keeps some of those gains and 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 and using (3), we obtain

$$q_1^* = q_N^* + \left(\frac{\alpha b^2}{2c - \alpha b^2}\right) \frac{t}{c}$$

$$= q_N^* + \frac{b}{c} \Delta i.$$

It is also easy to see from (8) 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 1 highlights the effects of the PTA. 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, because $S$’s marginal cost is lower than $p_w$. They represent trade diversion-induced 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. 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.

Interestingly, the PTA can lead to too much investment relative to the efficient level. Recall that without the agreement $i_N^* < i^e$ for any $\alpha \in [0, 1)$. Such an unambiguous ordering does not
exist under the PTA. Comparing $i^*_P$ with $i^e$,\textsuperscript{11} one finds that

$$i^*_P > i^e \iff (2c - b^2)\alpha t > 2c(1 - \alpha)(p_w - \omega).$$

It follows that $i^*_P > i^e$ when $\alpha$ is sufficiently close to one (in which case the 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 relationship" tradeoff between classic Vinerian trade diversion and an entirely novel effect. Due to the PTA, the firms create additional surplus for all units of customized inputs that would be produced without the PTA, plus some surplus for additional units traded. This surplus—areas $C$ and $D$ in Figure 1—increases welfare, possibly by more than the extra costs due to excessive production (area $E$) and investment.

It is important to stress at this point that, while our model displays 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 fixed by design (for given $t$). Thus, if one considered only tra-

\textsuperscript{11} In the Appendix we show that the efficient level of investment is the same under no agreement and under a PTA.
ditional forces, one would deem the model designed to highlight the negative welfare consequences of PTAs. However, in our environment we uncover a novel channel through which PTAs can raise economic efficiency.

5 Welfare Effects of a PTA

Consumer welfare from the final good remains constant, because the total level of inputs purchased, $Q^*$, does not change. But the 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 $S$.

We split the welfare impact of the PTA ($\Delta \Psi$) into the two effects, relationship strengthening ($\Delta \Psi_R$) and sourcing diversion ($\Delta \Psi_S$), so that $\Delta \Psi = \Delta \Psi_R + \Delta \Psi_S$. The relationship-strengthening effect reflects the welfare consequences of the PTA on the (ex-ante) investment decisions assuming that, given the investment, the (ex-post) sourcing decision is 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 analogous cost from using generic inputs in the production of the ex-post efficient level $q_1^*$—net of the increased investment cost. Specifically,

$$\Delta \Psi_R = p_w(q_1^* - q_N^*) + [C(q_N^*, i_N^*) - C(q_1^*, i_P^*)] - [I(i_P^*) - I(i_N^*)].$$ (11)

This expression can be rewritten as

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

$$= \Delta i \left[(1 - \alpha) bq_N^* + \left(\frac{b^2}{2c} - 1\right) \Delta i\right].$$ (12)

The first term inside the brackets is positive and proportional to the initial level of output. It reflects the welfare gains from moving the supplier’s output toward the first-best level, and is greater when the hold-up problem is more severe (that is, when the supplier has less bargaining power). The second term is negative and proportional to the cost of the investment.

In turn, the sourcing-diversion effect reflects the welfare consequences of the PTA on the (ex-
post) sourcing decisions, given the investment choice under the PTA. Basically, it corresponds to the deadweight loss from using customized inputs that are too expensive. This is the direct result of the protection the tariff preference effectively affords $S$ by skewing the sourcing decision away from the initial equilibrium. Explicitly,

$$
\Delta \Psi_S = C(q_1^*, i_P^*) - C(q_P, i_P) + \pi_w(q_P - q_1^*) \\
= -\frac{t^2}{2c}.
$$

(13)

This corresponds to (the negative of) area $E$ from Figure 1—a triangle with base $(q_P - q_1) = \frac{t}{c}$ and height $t$.

Looking first at the relationship-strengthening effect, we take the partial derivative of (12) with respect to the heterogeneity parameter $\omega$, and find

$$
\frac{\partial \Delta \Psi_R}{\partial \omega} = (1 - \alpha) b \Delta i \frac{\partial q_N^*}{\partial \omega} < 0.
$$

(14)

This expression is negative because $q_N^*$ is decreasing in $\omega$. It implies that the cost saving aspect of a PTA is unambiguously more important for more productive firms (which have a lower $\omega$). The key force behind this result is that more productive firms produce more (for given investment), and therefore the beneficial effect of the extra investment due to the PTA applies to more units than it would, were the supplier less productive.\(^{12}\)

In terms of Figure 1, this means that areas $C$ and $D$ expand as $\omega$ falls. Observe that the size of the vertical shift of the marginal curve is independent of $\omega$, since $\Delta i$ is unaffected by $\omega$. Similarly, the horizontal shift, given by $q_1 - q_N^*$, is also independent of $\omega$. It follows that the combined area $C + D$ is affected by $\omega$ only because it is proportional to $q_N^*$. Intuitively, a more productive supplier produces a larger number of units in equilibrium, with or without the PTA. Thus, when cost-reducing investment rises with the PTA, the corresponding cost savings apply to more units when $\omega$ is relatively low.

On the other hand, the sourcing-diversion effect does not change with $\omega$. Since neither the level of productivity nor investment affects the slope of the marginal cost curve, the implied deadweight

\(^{12}\) A similar result can be obtained if we look at other parameters that could index productivity in our model, such as $c$ (the slope of the marginal cost curve) or $b$ (the effectiveness of investment in lowering marginal costs).
loss is a constant function of both. The upshot is that the downside of an agreement is unaffected by the productivity of the supplier, whereas the upside rises with it.

**Lemma 1** As $S$’s productivity rises, the cost savings from a Preferential Trade Agreement unambiguously increase, but its sourcing diversion remains unchanged.

A central element behind Lemma 1 is that only the slope (and not the level) of the marginal cost curve affects the sourcing diversion effect. Since productivity only shifts vertically that curve, productivity does not influence the extent of sourcing diversion.\(^{13}\)

One implication of Lemma 1 is that the desirability of the inclusion of specific industries in the PTA is not all the same. We find that $\Delta \Psi_R + \Delta \Psi_S \geq 0$ if

$$\omega \leq p_w - \left[ \frac{2c - 2ab^2 + \alpha^2 b^2}{2\alpha(1 - \alpha)b^2} \right] t = \hat{\omega}$$  (15)

Observe that, since $2c > b^2$, the expression is brackets in strictly positive. It follows that, especially if $t$ is sufficiently high, $\hat{\omega}$ may be negative, in which case there are no suppliers in our model such that the PTA enhances welfare. However, that does not need to be the case:

**Proposition 1** The PTA enhances welfare only if $S$ is sufficiently productive: $\omega < \hat{\omega}$.

The level of the external tariff is an important element shaping the productivity threshold defined in Proposition 1. The effect is unambiguous: a higher $t$ yields a lower $\hat{\omega}$. In words: the more discriminatory is the PTA, the stricter is the supplier’s productivity requirement to make the agreement welfare improving. See the Appendix for the proof.

**Proposition 2** As the external tariff rises, the threshold level of productivity $\hat{\omega}$ unambiguously decreases, implying that more productive suppliers of customized inputs are required to ensure that the PTA will yield a net positive welfare effect.

At $\omega = \hat{\omega}$, the relationship-strengthening effect exactly balances the sourcing-diversion effect. At that point, an increase in $t$ does not improve the former as much as it worsens the latter. As a

\(^{13}\)Clearly, if marginal cost were not linear in $q$, Lemma 1 would no longer hold in its current simple form. For example, if marginal cost were convex in $q$ (and $\omega$ were still a horizontal shifter of the curve), the sourcing diversion effect would be smaller for high-productivity suppliers, as they would operate in a steeper portion of the marginal cost curve. This would reinforce the points we make below. More generally, a sufficient condition for the forthcoming conclusions to remain valid if marginal cost is not too concave in $q$. 

---

17
result, a decrease in $\omega$ is required to offset that difference. The reason is that the welfare impact of the PTA increases with supplier productivity (i.e., falls with $\omega$).

It is interesting to observe that in our setting, when it comes to evaluating the consequences of the PTA, society’s and producers’ views are aligned. Specifically, just as it happens with the two countries’ joint welfare, $\Psi(\omega)$, a PTA increases the payoff of the producers by more, the lower is $\omega$.

**Proposition 3** The PTA enhances the firms’ bargaining surplus by more, the higher the productivity of the supplier (i.e., the lower is $\omega$). The same is true for the firms’ total payoff.

**Proof.** The bargaining surplus of a pair $B-S$ changes with the PTA according to

$$\Delta \Omega(\omega) = p_w(q_P^* - q_N^*) + tq_P^* + [C(q_N^*, i_N^*) - C(q_P^*, i_P^*)].$$

Using (11), (13) and the definition of $I(i)$, this expression can be rewritten as

$$\Delta \Omega(\omega) = \Delta \Psi(\omega) + tq_P^* + \Delta i (i_P^* + i_N^*).$$

From the discussion leading to Proposition 1, we know that $\partial \Delta \Psi(\omega)/\partial \omega < 0$. We know from (9) and (10) that $\partial \Delta q_P^*/\partial \omega < 0$. And we know from (5) and from (9) that $\partial (i_P^* + i_N^*)/\partial \omega < 0$ as well. Hence, $\partial \Delta \Omega(\omega)/\partial \omega < 0$.

In turn, the firms’ total payoff changes with the PTA according to

$$\Delta \Omega(\omega) - [I (i_P^*) - I (i_N^*)] = \Delta \Psi(\omega) + tq_P^*.$$

Since $\partial \Delta \Psi(\omega)/\partial \omega < 0$ and $\partial \Delta q_P^*/\partial \omega < 0$, this expression falls with $\omega$. ■

Proposition 1 provides a criterion for industry exclusion in PTAs. Such “exceptions” are pervasive in reality, but there is very little research on the desirability of excluding industries when forming PTAs. An important exception is Grossman and Helpman (1995), but their analysis is positive rather than normative: because of producers’ lobbying pressures, governments tend to leave out precisely the efficiency-enhancing industries, while keeping those that would amplify inefficiencies. Our normative analysis suggests the opposite. This may seem unsurprising, but observe that the key players in our analysis are the producers. Put differently, in our setting the interests
of producers and of society are aligned, in the sense that Proposition 3 is analogous to Proposition 1, and they call for prioritizing industries with high underlying productivity in the agreement, contrarily to Grossman and Helpman’s conclusion. The key reason for the difference is that here a PTA has the benefit of lowering inefficiencies due to hold-up problems, not the focus of Grossman and Helpman (1995).

5.1 The Effect of the Tariff Preference

Rearranging (15), we see that the welfare effect of the PTA is positive if

\[ t \leq \left[ \frac{2\alpha(1 - \alpha)b^2}{2c - 2\alpha b^2 + \alpha^2 b^2} \right] (p_w - \omega) \equiv \hat{t}. \tag{16} \]

Unlike higher productivity, which has unambiguously positive welfare effects (and which monotonically increases \( \hat{t} \)), an increase in the tariff when \( t \in [0, \hat{t}] \) may either enhance or weaken the welfare impact of a PTA. Note first that \( t \) becomes the \textit{external} tariff under the preferential agreement. Its size affects the relationship-strengthening and the sourcing-diversion effects differently.\(^{15}\) We have the following result (see the Appendix for the proof):

**Proposition 4** The welfare effect of the PTA is strictly positive for \( t \in [0, \hat{t}] \) and has an inverted-U shape with respect to the external tariff, being maximized at \( t = \frac{\hat{t}}{2} \), where \( \hat{t} \) is defined in equation (16).

Sourcing diversion is a very simple function of the tariff—trade diversion monotonically increases with \( t \), and at an increasing rate. On the other hand, the relationship-strengthening effect is more nuanced. It is positive for sufficiently low \( t \), initially rises, but eventually falls with \( t \). This can be easily seen in equation (12), where the first term inside the brackets is positive and not a function of \( t \), the second term is negative and proportional to \( t \), and the whole brackets is multiplied by \( t \).

\(^{14}\)The alignment is not perfect because the producers do not internalize the effect of the PTA on tariff revenue. However, this additional effect only reinforces the main point of Proposition 1. The reason is that tariff revenue falls with the PTA by more, the lower \( \omega \) is. Therefore, the PTA benefit for the producers decreases even faster with \( \omega \) than the benefit for society. In fact, from the producers’ point of view there is a strictly positive benefit with the PTA for \textit{any} \( \omega \).

\(^{15}\)Naturally, the tariff affects welfare also through the conventional mechanism of inefficiently lowering the total volume of traded inputs, \( Q^* \). However, observe that under dual sourcing with and without the PTA, that effect is unchanged by the agreement.
For very low $t$, the sourcing-diversion effect is second-order small, so the first-order relationship-strengthening effect dominates even for relatively unproductive firms (i.e., $\omega$ close to $p_w$). But because the tariff is small, $\Delta i$ is small, so $\Delta \Psi_R$ is small and the effects of the PTA on welfare are minor. As $t$ increases, $\Delta i$ increases. For low levels of $t$ (i.e., $t < \frac{\hat{t}}{2}$), where the term in brackets in (12) is positive and big, the welfare gain from a PTA rises with $t$. For sufficiently high $t$ (i.e., $t > \frac{\hat{t}}{2}$), however, the increase in $\Delta \Psi_R$ is more than offset by a fall in $\Delta \Psi_S$, so the welfare improvement from a PTA falls with $t$. For $t > \hat{t}$, we have $\Delta \Psi < 0$.

5.2 The Effect of the Intensity of the Hold-Up Problem

Naturally, the welfare effect of the PTA depends crucially upon the intensity of the hold-up problem. The agreement may enhance overall welfare provided that it primarily serves to substitute for incomplete contracts for sufficiently productive firms. When the hold-up problem is very mild ($\alpha$ is near 1), there is no important contractual inefficiency to substitute for. In that case, a PTA distorts sourcing decisions and induces excessive relationship-specific investment. In terms of equation (12), observe that when $\alpha \to 1$, the first term inside the brackets approaches zero and the relationship-strengthening effect is negative for any tariff and any $\omega$. Thus, when $\alpha \to 1$, the tariff discrimination under the PTA brings only negative welfare effects.

Conversely, when the hold-up problem approaches its most serious level ($\alpha$ is near 0), the PTA is a poor substitute for incomplete contracts because the investment response to the PTA is too weak. In that case, a PTA merely distorts sourcing decisions. This is clear from (12), since $\lim_{\alpha \to 0} \Delta i = 0$. Thus, also when $\alpha \to 0$, the tariff discrimination under the PTA only brings undesirable effects.\footnote{It is easy to check that $\lim_{\alpha \to 1} \Delta \Psi(\alpha) < \lim_{\alpha \to 0} \Delta \Psi(\alpha) < 0$.}

It is when the supplier’s bargaining power is neither too low nor too high that PTAs can be effective. In that case, there is underinvestment but investment is responsive to the tariff discrimination engendered by a PTA. Thus, if the supplier is sufficiently productive (Proposition 1) and the hold-up problem is "moderate," a PTA can enhance social welfare even in a setting where classic trade creation is absent.

More precisely, there is a level of bargaining power that maximizes $\Delta \Psi$:

$$\alpha^O \equiv \frac{(p_w - \omega) 2c}{(p_w - \omega) (4c - b^2) + t(2c - b^2)},$$

(17)

\footnote{It is easy to check that $\lim_{\alpha \to 1} \Delta \Psi(\alpha) < \lim_{\alpha \to 0} \Delta \Psi(\alpha) < 0$.}
This value is decreasing in the external tariff. For example, as $t \to 0$, the value of $\alpha$ that maximizes $\Delta \Psi$ becomes 

$$\lim_{t \to 0} \alpha^O = \frac{2c}{4c - b^2},$$

which is bigger than 0.5. As the external tariff rises, $\alpha^O$ falls and eventually drops below 0.5. Intuitively, all else equal, a high external tariff implies a high $\Delta i$. But a strong investment boost is welfare improving if there is a severe hold-up problem to start with (that is, if $\alpha$ is small). Conversely, when the hold-up problem is mild (i.e., $\alpha$ is high), the PTA helps only if the external tariff is relatively small, limiting sourcing diversion and preventing excessive investment.

Summarizing this discussion:

**Proposition 5** The welfare effect of the PTA is strictly negative for both $\alpha \to 0$ and $\alpha \to 1$, and has an inverted-U shape with respect to the level of the supplier’s bargaining power, being maximized at $\alpha = \alpha^O$, where $\alpha^O$ is defined in equation (17). Furthermore, $\alpha^O$ is decreasing in $t$.

### 6 Search and Matching

Up to now, we have analyzed the consequences of the PTA for a given buyer-seller match. We now take a step back and consider the search and matching problem.

Let there exist a unit mass of buyers, but let the number of buyers be small relative to the number of suppliers. Specifically, if some measure of buyers is already matched, the distribution of remaining suppliers remains unchanged for the other buyers looking to match. Let the distribution of suppliers vary according to productivity $\omega$ on $[0, \bar{\omega}]$ according to distribution $G(\omega)$.

The search framework follows a standard "no recall" setup. The buyer’s goal is to find a supplier with high enough productivity. It costs $K > 0$ to attempt a single search, which yields one match. If a buyer likes his first match, then he and the supplier specialize towards each other and pursue investment and production according to the model analyzed in sections 2-5. If a buyer does not like his first match, he may pay another $K$ and search again. Indeed, he may search as many times as he likes. Because a failed search leaves the buyer with $K$ in sunk costs but otherwise right where he started, the choice to initiate an additional search does not depend on the number of previous searches.
Without a PTA, a buyer matched with a supplier with productivity \( \omega \) obtains a payoff of

\[
U_B^N(\omega) = [V(Q^*) - (p_w + t)Q^*] + (1 - \alpha)\Omega_N^*(\omega),
\]

where \( \Omega_N^*(\omega) \) is the bargaining surplus without a PTA (as defined in equation (4)) evaluated at the equilibrium levels of investment and production. Clearly, \( \partial \Omega_N^*(\omega) / \partial \omega < 0 \). On the other hand, the term in brackets is independent of \( \omega \).

Now denote the payoff of a buyer who has just found a supplier with productivity \( \omega' \) by \( V_B^N(\omega') \). We have that

\[
V_B^N(\omega') = \max \left\{ U_B^N(\omega'), \int_0^{\omega'} V_B^N(\omega)dG(\omega) - K \right\}.
\]

In words, \( B \)'s payoff is the highest between what he would obtain by keeping the match, \( U_B^N(\omega') \), and what he would obtain by searching again (net of search costs), \( \int_0^{\omega'} V_B^N(\omega)d\omega - K \).

The search equilibrium is characterized by a cutoff rule, where all buyers search until they find a supplier with sufficiently high productivity, \( \omega \leq \bar{\omega}_N \). The cutoff value \( \bar{\omega}_N \) is determined by the following condition:

\[
U_B^N(\bar{\omega}_N) = \int_0^{\bar{\omega}_N} V_B^N(\omega)dG(\omega) - K.
\]

Under this cutoff rule, if \( \omega > \bar{\omega}_N \), \( B \) searches again. In that case, using (20), we have that

\[
V_B^N(\omega) = \int_0^{\omega} V_B^N(\omega)dG(\omega) - K = U_B^N(\bar{\omega}_N). \]

If instead \( \omega \leq \bar{\omega}_N \), \( B \) stops searching, so \( V_B^N(\omega) = U_B^N(\omega) \). Substituting those expressions back into (20) and rearranging, we obtain

\[
U_B^N(\bar{\omega}_N) = \int_0^{\bar{\omega}_N} U_B^N(\omega)dG(\omega) + [1 - G(\bar{\omega}_N)]U_B^N(\bar{\omega}_N) - K.
\]

The left-hand side is the buyer’s payoff when he matches with a supplier with productivity \( \bar{\omega}_N \). The right-hand side is the expected payoff of the buyer when he searches, net of the search cost. Note that with probability \( 1 - G(\bar{\omega}) \) the buyer’s search fails and his expected future payoff (net of sunk costs) is \( U_B^N(\bar{\omega}_N) \). Rearranging (21), we find

\[
U_B^N(\bar{\omega}_N) = E[U_B^N(\omega|\omega \leq \bar{\omega}_N)] - \frac{K}{G(\bar{\omega}_N)},
\]

where the right-hand side can now be interpreted as the expected gross payoff of a buyer who finds
a successful match minus the total expected future cost of searching again.

The introduction of a PTA with Foreign affects the search equilibrium. The effect will depend on whether we consider a buyer matched in Foreign or in ROW. Let us start by considering a buyer who was matched in Foreign before the PTA.

### 6.1 The Effect of a PTA on Matches in Foreign

When a PTA is implemented between Home and Foreign, trade of specialized inputs between them no longer incur a tariff. As discussed earlier, by itself this raises the bargaining surplus between Home buyers and specialized suppliers from Foreign. In turn, the prospect of greater surplus induces suppliers to increase their relationship-specific investment. For both reasons, the payoff of buyers and suppliers in any given match increases with the PTA.

Furthermore, as Proposition 3 shows, the PTA boosts the firms’ payoffs by more, the more productive is the supplier. An implication is that the buyers’ incentives to search for suppliers in Foreign change with PTA. Since search costs are unaltered but the returns to a match increase disproportionately with more productive suppliers, buyers will be more demanding in their search under the PTA. As a result, the cutoff level of \( \omega \) under the PTA, \( \tilde{\omega}_P \), is lower than \( \tilde{\omega}_N \).

**Proposition 6** Under the PTA, the search cutoff in Foreign drops to \( \tilde{\omega}_P \), where \( \tilde{\omega}_P < \tilde{\omega}_N \). Therefore, buyers matched with suppliers in Foreign with productivity \( \omega \in (\tilde{\omega}_P, \tilde{\omega}_N] \) will dissolve their original partnerships and search again until they find a supplier with \( \omega \leq \tilde{\omega}_P \).

**Proof.** In the absence of a PTA, the net gain from keep searching when \( B \) has found a supplier with productivity \( \tilde{\omega}_N \) is \( \int_{0}^{\tilde{\omega}_N} U_B^N(\omega) dG(\omega) - K - G(\tilde{\omega}_N)U_B^N(\tilde{\omega}_N) \), which equals zero by construction. With the PTA, the net gain from keep searching in that case is instead \( \int_{0}^{\tilde{\omega}_N} U_B^P(\omega) dG(\omega) - K - G(\tilde{\omega}_N)U_B^P(\tilde{\omega}_N) \). The threshold characterizing optimal search under the PTA, \( \tilde{\omega}_P \), is lower than \( \tilde{\omega}_N \) if the gain from keep searching with a match \( \tilde{\omega}_N \) increases with the agreement. This happens if

\[
\int_{0}^{\tilde{\omega}_N} [U_B^P(\omega) - U_B^N(\omega)] dG(\omega) - \int_{0}^{\tilde{\omega}_N} [U_B^P(\tilde{\omega}_N) - U_B^N(\tilde{\omega}_N)] dG(\omega) > 0,
\]

or equivalently, if

\[
(1 - \alpha) \int_{0}^{\tilde{\omega}_N} [\Delta\Omega(\omega) - \Delta\Omega(\tilde{\omega}_N)] dG(\omega) > 0.
\]
But from Proposition 3 we know that \( \partial \Delta \Omega(\omega)/\partial \omega < 0 \). Thus, the term in brackets is strictly positive for all \( \omega < \bar{\omega}_N \), confirming that the inequality above holds.

Furthermore, the search cutoff under the PTA is lower (and therefore the extent of rematches greater), the higher the external tariff. The reason is that a higher external tariff enhances the buyer’s payoff under the PTA, but the increase is greater, the more productive is the supplier.

**Proposition 7** The search cutoff in Foreign, \( \bar{\omega}_P \), is decreasing in the external tariff.

**Proof.** Under the PTA, the threshold characterizing optimal search under the PTA, \( \bar{\omega}_P \), is defined implicitly by the equation

\[
\int_0^{\bar{\omega}_P} \left[ U_B^P(\bar{\omega}_P) - U_B^P(\omega) \right] d\omega + K = 0. \tag{23}
\]

To see how the external tariff affects \( \bar{\omega}_P \), we can use the implicit function theorem: \( d\bar{\omega}_P/dt = -\frac{\partial (\text{lhs}(23))/\partial t}{\partial (\text{lhs}(23))/\partial \bar{\omega}_P} \). Analogously to (18), the buyer’s payoff under the PTA is

\[
U_B^P(\omega) = [V(Q^*) - (p_w + t)Q^*] + (1 - \alpha)\Omega^*_P(\omega).
\]

Therefore,

\[
\frac{dU_B^P(\omega)}{dt} = \left( -Q^* - t \frac{dQ^*}{dt} \right) + (1 - \alpha) \left[ q_P^* + (p_w + t) \frac{dq_P^*}{dt} - C_q \frac{dq_P^*}{dt} - C_i \frac{di_P^*}{dt} \right] \\
= \left( -Q^* - t \frac{dQ^*}{dt} \right) + (1 - \alpha) \left[ q_P^* - C_i \frac{di_P^*}{dt} \right],
\]

where, from the first to the second line, we use the condition for privately optimal search under a PTA. Noticing that \( di_P^*/dt > 0 \) is independent of \( \omega \), and since \( C_i = -bq \), we then have that

\[
\frac{\partial (\text{lhs}(23))}{\partial t} = (1 - \alpha) \left( 1 + b \frac{di_P^*}{dt} \right) \int_0^{\bar{\omega}_P} [q_P^*(\bar{\omega}_P) - q_P^*(\omega)] d\omega < 0,
\]

where the inequality follows because \( q_P^*(\omega_P) \) is decreasing in \( \omega \). Now, using the Leibniz rule, we
have that
\[
\frac{\partial (\text{lhs}(23))}{\partial \tilde{\omega}_{P}} = U_{B}^{P}(\tilde{\omega}_{P})g(\tilde{\omega}_{P}) + \frac{dU_{B}^{P}(\tilde{\omega}_{P})}{d\tilde{\omega}_{P}}G(\tilde{\omega}_{P}) - U_{B}^{P}(\tilde{\omega}_{P})g(\tilde{\omega}_{P})
\]
\[
= \frac{dU_{B}^{P}(\tilde{\omega}_{P})}{d\tilde{\omega}_{P}}G(\tilde{\omega}_{P}) < 0.
\]

Therefore, \(d\tilde{\omega}_{P}/dt < 0\). ■

6.2 The Effect of a PTA on Matches in ROW

In the absence of a PTA, buyers from *Home* use the same cutoff rule when searching in *Foreign* and *ROW*. After *Home* forms the PTA with *Foreign*, however, a buyer with a match in *ROW* may want to forsake his current match and search now in *Foreign*. The incentives to do that are 2-fold. First, even if \(B\) were to match in *Foreign* with a supplier with the same level of productivity as he currently has, he would enjoy a benefit from the direct (no tariff on specialized inputs) and indirect (higher supplier investment) benefits from the PTA. Furthermore, because those gains are disproportionately higher for more productive suppliers, there is an additional benefit from searching in *Foreign* after the PTA, as the Proposition 6 shows.

As a result, there will be a new cutoff level \(\tilde{\omega}_{P}^{ROW}\), such that \(\tilde{\omega}_{P}^{ROW} < \tilde{\omega}_{P}\), describing the search equilibrium in *ROW*. All buyers initially matched in *ROW* with types \(\omega > \tilde{\omega}_{P}^{ROW}\) will seek new partners in *Foreign* after the PTA, according to the cutoff level \(\tilde{\omega}_{P}\).

**Proposition 8** Under the PTA, the search cutoff in *ROW* drops to \(\tilde{\omega}_{P}^{ROW}\), \(\tilde{\omega}_{P}^{ROW} < \tilde{\omega}_{P}\). Buyers matched with suppliers in *ROW* with productivity \(\omega \in (\tilde{\omega}_{P}^{ROW}, \tilde{\omega}_{N}]\) dissolve their original partnerships and search again until they find a supplier with \(\omega \leq \tilde{\omega}_{P}\) in *Foreign*. The cutoff \(\tilde{\omega}_{P}^{ROW}\) is decreasing in external tariff. It may be below zero, in which case all buyers leave *ROW* to *Foreign* with the PTA.

**Proof.** We know that a buyer with a match \(\tilde{\omega}_{P}\) in *ROW* does not have an incentive to search there again, since \(\tilde{\omega}_{P} < \tilde{\omega}_{N}\). Now, if a buyer with match \(\omega'\) wants to leave *ROW* to search in *Foreign*, then all buyers with matches \(\omega'' > \omega'\) want to do the same. Thus, it suffices to show that a buyer with a match \(\tilde{\omega}_{P}\) in *ROW* has an incentive to search in *Foreign*. 

25
In *ROW*, a buyer with a match $\tilde{\omega}_P$ has payoff $U_B^N(\tilde{\omega}_P)$. If he searches in *Foreign*, he obtains the expected value from searching there under the PTA, which by definition of the cutoff is equal to $U_B^P(\tilde{\omega}_P)$. But we know that $U_B^P(\tilde{\omega}_P) > U_B^N(\tilde{\omega}_P)$. Thus, a buyer with a match $\tilde{\omega}_P$ in *ROW* has an incentive to search in *Foreign*. There will be a lower cutoff $\tilde{\omega}_P^{ROW}$ such that $U_B^P(\tilde{\omega}_P) = U_B^N(\tilde{\omega}_P^{ROW})$. Observe that the external tariff affects the equation defining $\tilde{\omega}_P^{ROW}$, $U_B^P(\tilde{\omega}_P) - U_B^N(\tilde{\omega}_P^{ROW}) = 0$, only through its first term: we know from Proposition 7 that $\tilde{\omega}_P$ needs to fall [so that $U_B^N(\tilde{\omega}_P^{ROW})$ rises] to keep the equality satisfied.

For matches with suppliers with $\omega \leq \tilde{\omega}_P^{ROW}$, $B$ keeps his match in *ROW*. Nothing guarantees, however, that $\tilde{\omega}_P^{ROW} > 0$. ■

### 6.3 Welfare Effects of a PTA with Search and Matching

To assess the welfare consequences of a PTA with search and matching, it is useful to start with the limiting case where search costs are negligible.

**Proposition 9** As $K \to 0$, the PTA induces no additional search or re-matching, but is welfare-improving under the highest possible external tariff.

This is a trivial yet powerful result. It follows from the simple fact that, in the absence of search costs, buyers will search until they find a supplier with $\omega = 0$. Thus, when $K \to 0$, both $\tilde{\omega}_N \to 0$ and $\tilde{\omega}_P \to 0$. Hence, in that case the maximum external tariff that makes the PTA welfare-improving, $\hat{\omega}$, reaches its minimum level, $\left[ \frac{2\alpha(1-\alpha)b^2}{2c-2\alpha b^2 + \alpha^2 b^2} \right] p_w$.

An implication of this result is that, if the PTA lowers aggregate welfare, it is because either the external tariff or search costs—two policy variables that could potentially also be changed with the agreement—are too high.\(^{17}\)

Now, inefficiencies arise not only in investment decisions, but also in the extent of search efforts. Again, under some conditions a PTA can mitigate those inefficiencies.

**Proposition 10** In the absence of a trade agreement, there is too little search from a social standpoint.

\(^{17}\)In fact, there is evidence that external tariffs tend to fall upon the formation of free trade agreements, especially when they are initially high (see, in particular, Crivelli, 2016).
Proof. Without a PTA, $B$’s net gain from keeping searching when he has found a supplier with productivity $\tilde{\omega}_N$ is $\int_0^{\tilde{\omega}_N} U_B^N(\omega) dG(\omega) - K - G(\tilde{\omega}_N) U_B^N(\tilde{\omega}_N)$, which equals zero by construction. From a social standpoint, the net gain from keeping searching in that case is instead $\int_0^{\tilde{\omega}_N} \Psi(\omega) dG(\omega) - K - G(\tilde{\omega}_N) \Psi(\tilde{\omega}_N)$. If this expression is strictly positive, it implies that $B$ searches too little, from a social perspective. Using $\int_0^{\tilde{\omega}_N} U_B^N(\omega) dG(\omega) - K - G(\tilde{\omega}_N) U_B^N(\tilde{\omega}_N) = 0$, we have that this will happen if

$$\int_0^{\tilde{\omega}_N} \left[ \Psi(\omega) - U_B^N(\omega) \right] dG(\omega) - \int_0^{\tilde{\omega}_N} \left[ \Psi(\tilde{\omega}_N) - U_B^N(\tilde{\omega}_N) \right] dG(\omega) > 0. \quad (24)$$

Now observe that $\Psi(\omega) - U_B^N(\omega) = \alpha [p_w q_N^*(\omega) - C(q_N^*(\omega), i_N^*(\omega), \omega)] - I(i_N^*(\omega)) + tQ^*$. Hence, we can rewrite (24) as

$$\int_0^{\tilde{\omega}_N} \left\{ \alpha [p_w q_N^*(\omega) - C(q_N^*(\omega), i_N^*(\omega), \omega)] - I(i_N^*(\omega)) \right\} dG(\omega)$$

$$- \int_0^{\tilde{\omega}_N} \left\{ \alpha [p_w q_N^*(\tilde{\omega}_N) - C(q_N^*(\tilde{\omega}_N), i_N^*(\tilde{\omega}_N), \tilde{\omega}_N)] - I(i_N^*(\tilde{\omega}_N)) \right\} dG(\omega) > 0. \quad (25)$$

This inequality is satisfied, in particular, if the expression $\left\{ \alpha [p_w q_N^*(\omega) - C(q_N^*(\omega), i_N^*(\omega), \omega)] - I(i_N^*(\omega)) \right\}$ decreases with $\omega$. We have that

$$d \left\{ \alpha [p_w q_N^*(\omega) - C(q_N^*(\omega), i_N^*(\omega), \omega)] - I(i_N^*(\omega)) \right\}$$

$$= \frac{d\omega}{d\omega} \alpha \frac{dq_N^*}{d\omega} [p_w - C_q(q_N^*, i_N^*, \omega)] - \alpha C_\omega (q_N^*, i_N^*, \omega) - [\alpha C_i (q_N^*, i_N^*, \omega) + I'(i_N^*)].$$

But note that, due to the privately optimal conditions for sourcing and for the investment decision, both square brackets in this equation are nil. Therefore, $d\left\{ \alpha [p_w q_N^*(\omega) - C(q_N^*(\omega), i_N^*(\omega), \omega)] - I(i_N^*(\omega)) \right\}/d\omega$ reduces to $-\alpha C_\omega (q_N^*, i_N^*, \omega) = -\alpha q_N^* < 0$. It follows that inequality (25) holds and there is too little search from a social standpoint in the absence of a PTA. ■

Intuitively, the reason for this inefficiency is as follows. The buyer disregards $\alpha\%$ of the bargaining surplus as well as the supplier’s investment cost.\(^\text{18}\) Since $[\alpha \Omega(\cdot) - I(\cdot)]$ increases with productivity (decreases with $\omega$), the buyer searches too little from a social perspective.\(^\text{19}\)

\(^\text{18}\)The buyer also disregards tariff revenue when making his search decision. Note, however, that in the absence of a PTA tariff revenue ($\text{TR}$) is fixed at $\text{TR} = tQ^*$ by design. Therefore, it makes no difference for the efficiency of $B$’s search decision.

\(^\text{19}\)Note that, when $\alpha \to 0$, there is no investment and $B$ keeps the whole surplus from the transaction; as a result, in the limit when $\alpha \to 0$ he acts like a social planer in the search process (observe that the left-hand side of (25) goes
Proposition 11 Under a PTA, search in Foreign if efficient if \( t = \tilde{t} \equiv \frac{\omega}{\alpha} (p_w - \omega) \). If \( t < \tilde{t} \), search is inefficiently low; if \( t > \tilde{t} \), search is inefficiently high in Foreign. Furthermore, \( \tilde{t} < \tilde{\tilde{t}} \).

Proof. Under a PTA, B’s net gain from keeping searching in Foreign when he has found a supplier with productivity \( \tilde{\omega}_P \) is \( \int_0^{\tilde{\omega}_P} U_B'(P) dG(\omega) - K - G(\tilde{\omega}_P)U_B'(\tilde{\omega}_P) \), which equals zero by construction. From a social standpoint, the net gain from keeping searching in that case is instead \( \int_0^{\tilde{\omega}_P} \Psi(\omega)dG(\omega) - K - G(\tilde{\omega}_P)\Psi(\tilde{\omega}_P) \). If this expression is strictly positive, it implies that search is too low from a social perspective; if it is strictly negative, it implies that search is inefficiently high. Using \( \int_0^{\tilde{\omega}_P} U_B'(P) dG(\omega) - K - G(\tilde{\omega}_P)U_B'(\tilde{\omega}_P) = 0 \), we have that search is inefficiently low if

\[
\int_0^{\tilde{\omega}_P} \left[ \Psi(\omega) - U_B'(\omega) \right] dG(\omega) - \int_0^{\tilde{\omega}_P} \left[ \Psi(\tilde{\omega}_P) - U_B'(\tilde{\omega}_P) \right] dG(\omega) > 0.
\]

Now observe that \( \Psi'(\omega) - U_B'(\omega) = \alpha [(p_w + t) q_P^* - C (q_P^*, i_P^*, \omega)] - I (i_P^*) + tQ^* - t q_P^* \). Hence, we can rewrite (6.3) as

\[
\begin{align*}
&\int_0^{\tilde{\omega}_P} \left\{ \alpha [(p_w + t) q_P^*(\omega) - C (q_P^*(\omega), i_P^*(\omega), \omega)] - I (i_P^*(\omega)) - t q_P^*(\omega) \right\} dG(\omega) \\
&\quad - \int_0^{\tilde{\omega}_P} \left\{ \alpha [(p_w + t) q_P^*(\tilde{\omega}_P) - C (q_P^*(\tilde{\omega}_P), i_P^*(\tilde{\omega}_P), \tilde{\omega}_P)] - I (i_P^*(\tilde{\omega}_P)) - t q_P^*(\tilde{\omega}_P) \right\} dG(\omega) > 0.
\end{align*}
\]

This inequality is satisfied, in particular, if the expression \( \left\{ \alpha [(p_w + t) q_P^*(\omega) - C (q_P^*(\omega), i_P^*(\omega), \omega)] - I (i_P^*(\omega)) - t q_P^*(\omega) \right\} \) decreases with \( \omega \). But we have that

\[
\frac{d \left\{ \alpha [(p_w + t) q_P^*(\omega) - C (q_P^*(\omega), i_P^*(\omega), \omega)] - I (i_P^*(\omega)) - t q_P^*(\omega) \right\}}{d\omega} = \alpha \frac{dq_P^*}{d\omega} [p_w + t - C_q (q_P^*, i_P^*, \omega)] - \alpha C_{i}\omega (q_P^*, i_P^*, \omega) - \left[ \alpha C_i (q_P^*, i_P^*, \omega) + I' (i_P^*) \right] \frac{di_P^*}{d\omega} - t \frac{dq_P^*}{d\omega}.
\]

Due to the privately optimal conditions for sourcing and for the investment decision, both square brackets in this equation are nil. Therefore, \( \frac{d \left\{ \alpha [(p_w + t) q_P^*(\omega) - C (q_P^*(\omega), i_P^*(\omega), \omega)] - I (i_P^*(\omega)) - t q_P^*(\omega) \right\}}{d\omega} \) reduces to \(- \alpha C_{\omega} (q_P^*, i_P^*, \omega) - t \frac{dq_P^*}{d\omega} \). Using equations (9) and (10), this expression can be rewritten as \( \frac{2}{\alpha - \alpha t} [(1 - \alpha)(t - \alpha(p_w - \omega))] \), which is negative if \( t < \frac{\alpha}{1 - \omega} (p_w - \omega) \equiv \tilde{t} \).

Thus, if \( t < \tilde{t} \), there is too little search from a social standpoint also under the PTA. Analogous reasoning implies that, if \( t > \tilde{t} \), there is excessive search from a social standpoint under the PTA.
If $t = \tilde{t}$, search is efficient under the PTA. Finally, observe from (16) that $\frac{i}{2} < \tilde{t}$ if

$$\left[ \frac{\alpha(1 - \alpha)b^2}{2c - 2\alpha b^2 + \alpha^2 b^2} \right] (p_w - \omega) < \frac{\alpha}{1 - \alpha} (p_w - \omega)$$

$$\iff (1 - \alpha)^2 b^2 < 2c - 2\alpha b^2 + \alpha^2 b^2$$

$$\iff [(1 - \alpha)^2 + 2\alpha - \alpha^2] b^2 < 2c$$

$$\iff b^2 < 2c.$$ 

which is always true. ■

Hence, although in the absence of a preferential trade agreement there is too little search, this may not happen under a PTA if the external tariff is sufficiently high, in which case the tariff preference would generate excessive search by the buyer and make the cutoff $\tilde{\omega}_P$ become too low. The reason is that, under a PTA, $B$ doesn’t internalize that searching for a lower $\omega$ implies greater loss of tariff revenue.

Still, if the external tariff is lower than the level that maximizes the welfare gains from the PTA absent search, $\frac{i}{2}$, then there would be too little search from a social perspective even under the PTA. Search will be efficient only under a specific level of $t$ that exactly balances the original buyer’s disincentive to search efficiently due to him not internalizing the full benefits of search and his excessive incentive to search due to the tariff preference.\textsuperscript{20}

\textsuperscript{20}Note that, when $\alpha \to 0$, the tariff threshold goes to zero. The reason is that, in such a case, search is not inefficient to start with, and as result the PTA induces excessive search for sure. Conversely, when $\alpha \to 1$, the tariff threshold goes to infinite. The reason is that, in that case, without the PTA the buyer has very little incentive to search. As a result, the PTA helps for sure in that dimension.
6.3.1 Welfare impact of a PTA: ‘comparative statics’

Observe that, with search:

- Expected payoff for $B$ without PTA: $U_B^N(\bar{\omega}_N)$.
- Expected payoff for $B$ with PTA: $U_B^P(\bar{\omega}_P)$.

From that, we conclude (?) that:

- Expected welfare without PTA: $\Psi^N(\bar{\omega}_N)$.
- Expected welfare with PTA: $\Psi^P(\bar{\omega}_P)$.

Welfare impact of PTA:

$$
\Delta \Psi^{search} = \Psi^P(\bar{\omega}_P) - \Psi^N(\bar{\omega}_N)
= \Psi^P(\bar{\omega}_P) - \Psi^P(\bar{\omega}_N) + \Psi^P(\bar{\omega}_N) - \Psi^N(\bar{\omega}_N)
= \Psi^P(\bar{\omega}_P) - \Psi^P(\bar{\omega}_N) + \Delta \Psi(\bar{\omega}_N).
$$

Thus, $\Delta \Psi^{search} - \Delta \Psi^{nosearch} = \Psi^P(\bar{\omega}_P) - \Psi^P(\bar{\omega}_N)$. Now observe that $\Psi^P(\omega) = U_{B-S}^P(\omega) + t [Q^* - q^*_P(\omega)]$, where $U_{B-S}^P$ denotes the joint payoff of the buyer and the seller in a relationship. Recalling that $U_{B-S}^P(\omega) = \Omega_P^*(\omega) - I(i_P^*(\omega)) = (p_w + t) q^*_P(\omega) - C(q^*_P(\omega), i_P^*(\omega)) - I(i_P^*(\omega))$, it follows that

$$
\Delta \Psi^{search} - \Delta \Psi^{nosearch} = [p_w q^*_P(\bar{\omega}_P) - C(q^*_P(\bar{\omega}_P), i_P^*(\bar{\omega}_P)) - I(i_P^*(\bar{\omega}_P))]
- [p_w q^*_P(\bar{\omega}_N) - C(q^*_P(\bar{\omega}_N), i_P^*(\bar{\omega}_N)) - I(i_P^*(\bar{\omega}_N))].
$$

We then have the following result.

**Proposition 12** The welfare impact of the PTA for matches in Foreign is always enhanced by B’s search effort.
Proof. The expression in (26) is positive when
\[
-\int_{\bar{\omega}}^{\tilde{\omega}} \left[ p_w q_P^*(\omega) - C(q_P^*(\omega), i^*_P(\omega)) - I(i^*_P(\omega)) \right] d\omega > 0.
\]

This inequality necessarily holds when the derivative inside the integral is negative. Now observe that, after using the condition for the investment decision \((I'(i^*_P) = -\alpha C_i(\cdot))\) and the condition for optimal private sourcing under the PTA (equation (7)), that derivative can be rewritten as
\[
\frac{d}{d\omega} \left[ p_w q_P^*(\omega) - C(q_P^*(\omega), i^*_P(\omega)) - I(i^*_P(\omega)) \right] = -t \frac{d q_P^*(\omega)}{d\omega} - (1 - \alpha) C_i(\cdot) \frac{d i_P^*(\omega)}{d\omega} - C(\cdot).
\]

The first term of (27) is positive, but the other two are negative. To see that the whole expression is always negative, it suffices to observe that
\[
-t \frac{d q_P^*(\omega)}{d\omega} - C(\cdot) = -t \frac{d q_P^*(\omega)}{d\omega} - q_P^*(\omega) = \frac{2t}{2c - \alpha b^2} - \frac{2(p_w + t - \omega)}{2c - \alpha b^2} = -2 \frac{p_w - \omega}{2c - \alpha b^2} < 0,
\]
concluding the proof. ■

This results follows from the fact that, for a given trade regime, higher supplier productivity necessarily enhances welfare. This is not obvious, because higher productivity yields a greater loss of tariff revenue under a PTA—the first term of equation (27). Nevertheless, as the proof above shows, that effect is more than offset by the direct effect of higher productivity in reducing the production cost of customized components—the third term of (27). Furthermore, higher productivity is also associated with greater supplier investment—the second term of (27). Hence, if a PTA raises welfare without search (which necessarily happens, for example, if \(\bar{\omega}_N \leq \tilde{\omega}\)), it raises welfare even more when search for suppliers is accounted for. Analogously, even if a PTA does not raise welfare in the absence of search, it may do so once search for suppliers is accounted for, provided that the difference \((\bar{\omega}_N - \tilde{\omega}_P)\) is high enough.
If the last proposition is right, we can delete the text below:

Considering search and matching, we can write welfare with and without the PTA as follows:

\[
W^N = \int_0^{\tilde{\omega}_N} \Psi(\omega, 0) dG(\omega) + [1 - G(\tilde{\omega}_N)] \Psi(\tilde{\omega}_N(0), 0) - K
\]

\[
W^P = \int_0^{\tilde{\omega}_P} \Psi(\omega, t) dG(\omega) + [1 - G(\tilde{\omega}_P)] \Psi(\tilde{\omega}_P(t), t) - K.
\]

Then the change in welfare induced by the PTA is

\[
\Delta W = \int_0^{\tilde{\omega}_P} \Delta \Psi(\omega) dG(\omega) + \int_0^{\tilde{\omega}_N} [\Psi(\tilde{\omega}_P(t), t) - \Psi(\omega, 0)] dG(\omega) + [1 - G(\tilde{\omega}_N)][\Psi(\tilde{\omega}_P(t), t) - \Psi(\tilde{\omega}_N(0), 0)].
\]

In general, this full effect is ambiguous, because the \(\Delta \Psi\) terms could all be negative if the tariff is sufficiently high. But compare this to the change in welfare without search,

\[
\Delta W_{\text{NoSearch}} = \int_0^{p_w} \Delta \Psi(\omega) dG(\omega),
\]

and we see that the difference-in-difference is

\[
\Delta \Delta W = \int_0^{\tilde{\omega}_N} \{\Psi(\tilde{\omega}_P(t), t) - \Psi(\omega, 0) - [\Psi(\omega, t) - \Psi(\omega, 0)]\} dG(\omega)
\]

\[
+ \int_0^{p_w} \{\Psi(\tilde{\omega}_P(t), t) - \Psi(\tilde{\omega}_N(0), 0) - [\Psi(\omega, t) - \Psi(\omega, 0)]\} dG(\omega)
\]

\[
= \int_0^{\tilde{\omega}_P} \{\Psi(\tilde{\omega}_P(t), t) - \Psi(\omega, t)\} dG(\omega)
\]

\[
+ \int_0^{p_w} \{\Psi(\tilde{\omega}_P(t), t) - \Psi(\tilde{\omega}_N(0), 0) - [\Psi(\omega, t) - \Psi(\omega, 0)]\} dG(\omega).
\]

Note that for the matches made both with and without a PTA \((\omega < \tilde{\omega}_P)\), there is no difference in the welfare impacts. The first term above is unambiguously positive, because \(\Psi\) is a decreasing function of \(\omega\). The second term is also positive whenever the PTA increases welfare for the cutoff supplier, because \(\Psi(\tilde{\omega}_P, t) - \Psi(\tilde{\omega}_P, 0)\) implies \(\Psi(\tilde{\omega}_P(t), t) - \Psi(\tilde{\omega}_N(0), 0) > 0\), and it is obvious that \(\Psi(\tilde{\omega}_P(t), t) - \Psi(\tilde{\omega}_N(0), 0) > \Delta \Psi(\omega)\) for all \(\omega \geq \tilde{\omega}_N\).
6.3.2 Welfare impact of a PTA: ‘MIT shock’

For matches initially in Foreign

- If the original match is \( \omega \leq \tilde{\omega}_P \), the PTA does not involve re-matching, so we are back to the benchmark case, except that the distribution of \( \omega \) is truncated at \( \tilde{\omega}_N \). Thus, expected welfare impact for those matches is \( \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} \).

- If the original match is \( \omega > \tilde{\omega}_P \), the PTA involves re-matching. In that case, net expected welfare generated by each \( B \) who rematches is \( \Psi^P(\tilde{\omega}_P) \). And again, the distribution of \( \omega \) is truncated at \( \tilde{\omega}_N \). Thus, expected welfare impact for those matches is \( \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^N(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)} \).

Hence,

\[
\Delta \Psi_{\text{search}}^* = \int_0^{\tilde{\omega}_P} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} + \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) - \Psi^N(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)}
\]

\[
= \int_0^{\tilde{\omega}_N} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} + \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) + \Delta \Psi(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)}
\]

\[
= \int_0^{\tilde{\omega}_N} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} + \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)}. \tag{28}
\]

The first term of (28) corresponds to the benchmark impact of the PTA, except that it is restricted to \( \omega \leq \tilde{\omega}_N \). The second term of (28) represents the welfare gain due to the endogenous increase in productivity for those with initially low-productivity matches, and is necessarily positive. Therefore,

\[
\Delta \Psi_{\text{search}}^* - \Delta \Psi_{\text{no-search}}^* = [1 - G(\tilde{\omega}_N)] \left[ \int_0^{\tilde{\omega}_N} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} - \int_{\tilde{\omega}_N}^{\tilde{\omega}_P} \Delta \Psi(\omega) \frac{dG(\omega)}{1 - G(\tilde{\omega}_N)} \right] + \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)},
\]

or equivalently,

\[
\Delta \Psi_{\text{search}}^* - \Delta \Psi_{\text{no-search}}^* = [1 - G(\tilde{\omega}_N)] \left\{ E[\Delta \Psi(\omega; \omega \leq \tilde{\omega}_N)] - E[\Delta \Psi(\omega; \omega \in [\tilde{\omega}_N, \tilde{\omega}_P]) + \int_{\tilde{\omega}_P}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)} \right\}.
\]

Since the term in curly brackets is positive because of the truncation of the original distribution of productivity, the whole expression is necessarily positive. First, regardless of the trade regime,
search induces buyers to seek for relatively good suppliers, which enhance the impact of the PTA. Second, a PTA induces buyers with relatively weak matches to search for more productive suppliers.

**Proposition 13** The additional welfare impact of the PTA due to possible rematches is necessarily positive for those buyers initially matched in Foreign.

**For matches initially in **ROW**

- If the original match is \( \omega \leq \tilde{\omega}_P^{ROW} \), the PTA does not involve re-matching. As those buyers remain matched outside the PTA bloc, the agreement is inconsequential for them.

- If the original match is \( \omega > \tilde{\omega}_P^{ROW} \), the PTA involves re-matching in Foreign. In that case, net expected welfare generated by each \( B \) who rematches is \( \Psi^P(\tilde{\omega}_P) \). As with the original matches in Foreign again, the distribution of \( \omega \) is truncated at \( \tilde{\omega}_N \). Thus, expected welfare impact for those matches is \( \int_{\tilde{\omega}_N}^{\tilde{\omega}_N^{ROW}} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^N(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)} \).

Hence,

\[
\Delta \Psi_{ROW}^{\text{search}} = \int_{\tilde{\omega}_N^{ROW}}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) + \Psi^P(\omega) - \Psi^N(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)} \\
= \int_{\tilde{\omega}_N^{ROW}}^{\tilde{\omega}_N} \Delta \Psi(\omega) \frac{dG(\omega)}{G(\tilde{\omega}_N)} + \int_{\tilde{\omega}_N^{ROW}}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)}. \tag{29}
\]

The first term of (29) corresponds to the benchmark impact of the PTA, restricted to \( \omega \in [\tilde{\omega}_P^{ROW}, \tilde{\omega}_N] \), as if those matches were originally in Foreign. The second term of (29) represents the welfare effect due to the endogenous change in productivity for those with initially low-productivity matches, again as if those matches were originally in Foreign. Unlike the corresponding expression for matches actually in Foreign before the PTA, that expression is not necessarily positive. The reason is that \( \tilde{\omega}_P > \tilde{\omega}_P^{ROW} \), so it is possible that some buyers leaving ROW end up with matches with lower productivity than they initially had.
In the absence of active search, those buyers initially in ROW would be unaffected by the PTA. Therefore,

\[ \Delta \Psi_{\text{search}}^{\text{ROW}} - \Delta \Psi_{\text{no-search}}^{\text{ROW}} = \]

\[ \Delta \Psi_{\text{search}}^{\text{ROW}} = E \left[ \Delta \Psi(\omega; \omega \in [\tilde{\omega}_P^{ROW}, \tilde{\omega}_N]) \right] + \int_{\tilde{\omega}_P^{ROW}}^{\tilde{\omega}_N} \left[ \Psi^P(\tilde{\omega}_P) - \Psi^P(\omega) \right] \frac{dG(\omega)}{G(\tilde{\omega}_N)}. \]

In contrast to the analogous expression for matches initially in Foreign, the sign of the expression is ambiguous. First, the benchmark welfare impact of the PTA for buyers with matches \( \omega \in [\tilde{\omega}_P^{ROW}, \tilde{\omega}_N] \) may be negative, and that negative effect would not have been felt if the buyers did not move from ROW. Second, the buyers who leave ROW to move to Foreign because of the PTA may end up with worse matches than they originally had.

**Proposition 14** The additional welfare impact of the PTA due to possible rematches is ambiguous for those buyers initially matched in ROW.

7 Conclusion

In this paper, we offer novel insights on the welfare effects of PTAs in the context of traded customized intermediate inputs in the presence of incomplete contracts. In that context, we show that a PTA stimulates relationship-specific investments that may generate enough additional surplus to overwhelm traditional Vinerian trade diversion effects. This will tend to occur among firms that already trade significantly, because it is for such firms that investments create the most surplus. Furthermore, a PTA induces firms to search more intensively for better partners. As this enhances the quality of the matches, this further enhances the gains with the agreement.

We obtain those results for the simple case of linear demand. This specification has numerous useful characteristics. Primarily, it sharply highlights the key first-order effect of the pre-PTA level of inputs on whether the agreement is likely to enhance welfare. The model permits simple closed-form solutions for welfare and shows that the positive force for welfare is proportional to the initial level of inputs. Alternative specifications—e.g., non-linear marginal costs—would introduce second-order effects that would blur this connection but not diminish the first-order importance of
the initial level of inputs. This specification also restricts attention to cases where a tariff causes the same trade diversion effect for all firms that produce customized inputs. This helps to fix ideas and sacrifices little generality. This effect is the product of the size of the tariff (which is the same for all firms by construction) and the number of units that are diverted. This latter figure would differ across firms if the marginal cost curve were non-linear. Still, those differences would be of second-order importance.

We impose dual sourcing to shut down classic trade creation in the analysis. This permits us to emphasize the trade-off between the costs from trade diversion and the benefits from efficiency-enhancing higher investment due to a PTA. But observe that, if the investment effect were sufficiently high, dual would become single sourcing, and the PTA would also generate Vinerian trade creation. This would tilt the trade-off toward higher net gains.

At a more general level, an increasingly important theme for policymakers and academics alike is the expansion of global value chains (GVCs). Our setting is rather simple, with a GVC containing only two firms and with inputs crossing only one national border, whereas a typical GVC includes several producers and parts crossing 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 ‘true’ GVCs.

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 show that preferential market access is critical for the international fragmentation of production, in fact probably more than it has ever been for the trade of final goods. Hence, we provide a step towards a framework that incorporates the Vinerian view into the “new regionalism.” Much more is needed, though.
Appendix

Efficient investment levels Without an agreement, the efficient investment level solves
\[
\max_i p_w q_N - C(q_N, i, \omega) - I(i). \tag{30}
\]
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 (2), this expression simplifies to \(-C_i(q_N, i, \omega) = I'(i^e)\), as indicated in (6).

With a PTA, the efficient investment level also solves (30), after replacing \(q_N\) with \(q_P\). The first-order necessary condition is analogous to the one above, but using (7) 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 \(i = i^e\).

Proof of Proposition 2. The productivity threshold \(\hat{\omega}\) is defined as the level of \(\omega\) that makes the PTA welfare-neutral. That is, \(\Delta \Psi(\hat{\omega}, t) \equiv 0\). Using (12) and (13), this expression can be rewritten as
\[
\Delta t \left[ (1 - \alpha)b q_N^*(\hat{\omega}) + \left( \frac{b^2}{2c} - 1 \right) \Delta i \right] - \frac{t^2}{2c} = 0,
\]
or equivalently,
\[
\left[ (1 - \alpha)b q_N^*(\hat{\omega}) + \left( \frac{b^2}{2c} - 1 \right) \Delta i \right] = \frac{t^2}{2c \Delta i}. \tag{31}
\]
The effect of the tariff on \(\hat{\omega}\) is given by
\[
\frac{d\hat{\omega}}{dt} = - \frac{\partial \Delta \Psi(\hat{\omega}, t)/\partial t}{\partial \Delta \Psi(\hat{\omega}, t)/\partial \omega}.
\]
We know from (14) that the denominator of this expression is strictly negative. It follows that 

\[ \text{sgn}(d\dot{\omega}/dt) = \text{sgn}(\partial \Delta \Psi(\dot{\omega}, t)/\partial t). \]

To determine \( \text{sgn}(\partial \Delta \Psi(\dot{\omega}, t)/\partial t) \), use (31) and recall that \( dq^*_{N}/dt = 0 \) to obtain

\[ \frac{\partial \Delta \Psi(\dot{\omega}, t)}{\partial t} = \frac{\partial \Delta i}{\partial t} \left[ \frac{t^2}{2c\Delta i} - \Delta i \left( \frac{2c - b^2}{2c} \right) \right] - \frac{t}{c}. \]

Now notice that, as \( \Delta i \) is a linear function of \( t \), \( \frac{\partial \Delta i}{\partial t} \frac{t}{\Delta i} = 1 \). Thus,

\[ \frac{\partial \Delta \Psi(\dot{\omega}, t)}{\partial t} = -\frac{t}{2c} - \Delta i \left( \frac{2c - b^2}{2c} \right) \frac{\partial \Delta i}{\partial t} < 0. \]

Therefore, \( d\dot{\omega}/dt < 0. \)

**Proof of Proposition 4.** From (12) and (13), after some manipulation, we obtain

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

where the first term in the right-hand side corresponds to \( d\Delta \Psi_R/dt \) and the second to \( d\Delta \Psi_S/dt \). Equating this equation to zero and solving for \( t \), we then find

\[ t = \left( \frac{\alpha(1 - \alpha)b^2}{2c - 2ab^2 + \alpha^2b^2} \right) (p_w - \omega) = \frac{i}{2}. \]

In turn, the second-order necessary condition is satisfied provided that \( 2c - 2ab^2 + \alpha^2b^2 > 0 \), which is true since \( 2c > b^2 \) and because \( \alpha(2 - \alpha) < 1. \)

**References**


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


