Offshoring and the Value of Trade Agreements*

Pol Antràs
Harvard Unversity and NBER

Robert W. Staiger
Stanford University and NBER

Preliminary Version. Comments Welcome.

February 6, 2008

Abstract

We study the trade policy choices of governments in an environment in which some of the trade flows being taxed or subsidized involve the exchange of customized inputs, and the contracts governing these transactions are incomplete. We show that the second-best policies that emerge in this environment entail free trade in final goods but *not* in intermediate inputs, since import or export subsidies targeted to inputs can alleviate the international hold-up problem. We next show that the Nash equilibrium policy choices of governments do not coincide with internationally efficient choices, and that the Nash policies imply an inefficiently low level of intermediate input trade across countries. The reason is that in our environment trade policy choices serve a dual role: they can enhance investment by suppliers but, because of ex-post bargaining over prices, they can also be used to redistribute profits across countries. The inefficiencies inherent in the Nash policy choices of governments not only result in suboptimal input subsidies, but also in positive distortions in final-good prices, even when countries cannot affect world (untaxed) prices in those goods. As a result, an international trade agreement that brings countries to the efficiency frontier will necessarily increase trade in inputs, but it may require a reduction in final-goods trade. When governments are not motivated by the impact of their policies on ex-post negotiated international input prices, the resulting policy choices are efficient, and hence a modified terms-of-trade interpretation of the purpose of trade agreements can be offered, but *only* when governments maximize real national income. If governments preferences are sensitive to political economy (distributional) concerns, the purpose of a trade agreement becomes more complex, and cannot be reduced to solving a simple terms-of-trade problem.

*We thank Kyle Bagwell, Meredith Crowley, Elhanan Helpman, Alan Sykes and seminar participants at Northwestern University and the NBER ITI Winter Meeting for helpful comments and discussions. Staiger gratefully acknowledges financial support from the NSF (SES-0518802).

1 Introduction

International trade in intermediate inputs is a dominant feature of the world economy. Using OECD input-output tables, Ramanarayanan (2006) concludes that in the late 1990s imports of intermediate goods comprised about forty to sixty percent of total merchandise imports for a large number of OECD countries. Similarly, a thorough examination of highly disaggregated trade data led Yeats (2001) to conclude that intermediate input trade accounted for roughly thirty percent of world trade in manufacturing goods in 1995. Furthermore, several authors have noted that the share of intermediate inputs in world trade appears to have increased significantly in recent years.[1]

Recent developments in international trade theory have attempted to bridge the apparent gap between the characteristics of international trade in the data and the standard representation of these trade flows in terms of final goods in neoclassical trade theory. One branch of this new literature has focused on incorporating input trade in otherwise standard models with perfectly competitive markets and frictionless contracting.[2] Another branch of the literature has stressed that modelling offshoring as simply an increase in the fragmentation of production across countries misses important aspects of the characteristics of intermediate input trade.[3] First, intermediate inputs tend to be much more customized to their intended buyers than final goods, and hence, input trade embodies a disproportionate amount of relationship-specific investments. Second, because contracts involving international transactions are especially hard to enforce, the cross-border exchange of specialized intermediate inputs cannot generally be governed by the same contractual safeguards that usually accompany similar exchanges occurring within borders. A third distinguishing feature of international offshoring is that it is often associated with the costly search for suitable suppliers that can provide the required inputs.

The purpose of this paper is to study the Nash equilibrium and internationally efficient trade policy choices of governments in an environment in which some of the trade flows being taxed or subsidized involve the exchange of customized inputs, contracts governing these transactions are incomplete, and the matching between final-good producers and suppliers may involve search frictions. By identifying circumstances under which the Nash and internationally efficient policies do not correspond to one another in this environment, we thereby seek to understand the value of trade agreements in the presence of offshoring.

The first result of the paper is that the second-best policies that emerge in this environment entail free trade in final goods but *not* in intermediate inputs. Intuitively, the combination of

[1] See, for instance, Feenstra and Hanson (1996b), Feenstra (1998), Campa and Goldberg (1997), Hummels, Ishii and Yi (2001) and Borga and Zeile (2004).

[2] See for instance the work of Feenstra and Hanson (1996a), Jones (2000), Deardorff (2001), Antràs, Garicano and Rossi-Hansberg (2006), or Grossman and Rossi-Hansberg (2006).

[3] Theoretical developments include the work of McLaren (2000), Grossman and Helpman (2002, 2005), Antràs (2003, 2005), and Antràs and Helpman (2004). The empirical work of Feenstra and Hanson (2005), Levchenko (2007) and Nunn (2007) substantiate the empirical relevance of these non-standard features of offshoring.

relationship-specific investments and incomplete contracts results in a hold-up problem that leads to an inefficiently low volume of input trade across countries, and import or export subsidies targeted to these inputs can help bring countries closer to the efficiency frontier. In fact, we show that by choosing these subsidies appropriately, the international hold-up problem can be fully resolved and the efficiency frontier or first-best can be attained.

We next tackle the question of whether the Nash equilibrium policy choices of governments will coincide with the internationally efficient policies. We answer this question in the negative, and we show that the Nash equilibrium in trade policies always involves an inefficiently low level of intermediate input trade across countries. The broad intuition for the result is simple. Trade policy serves a dual role in a world with incomplete contracting in vertical relationships. On the one hand, as mentioned above, *subsidies* to the exchange of intermediate inputs can serve as a substitute for more standard contractual safeguards available in domestic transactions and can thus increase the volume of input trade toward its efficient level. On the other hand, in a world where input prices are not fixed in an enforceable initial (or ex-ante) contract, we show that trade *taxes* can also be used to redistribute surplus across countries. For instance, although an export tax in the input-producing country may reduce the incentives to invest of suppliers, in their ex-post bargaining with their buyers, these suppliers will be able to pass part of the cost of the tax on to final-good producers abroad. There is hence a basic tension that each government faces in its unilateral trade policy choices between correcting the hold-up problem and capturing surplus from its trading partner, and this tension prevents governments from making internationally efficient policy choices in the absence of a trade agreement (i.e., in the Nash equilibrium).

In order to illustrate these forces in the simplest possible way, we develop a model with two countries and a single taxable homogenous final good produced with a customized input, which can also be taxed. We assume that the two countries are small in the sense that their choices of trade taxes have no effect on the world (untaxed) price of the final good. For simplicity, we assume that all final-good producers are located in one of the two countries ("Home"), while all input producers are located in the other country ("Foreign").[4]

In the absence of frictions in input markets, welfare-maximizing governments would have no incentive to tax final-good trade, and as mentioned above, second-best trade policies would also involve free trade in the final good. We show, however, that in the Nash equilibrium in trade policies, the input-importing country will generally have an incentive to tax or subsidize not only the intermediate input, but also the final good. This follows from the fact that by affecting the surplus and outside options of both producers, trade taxes on the final good can also serve a surplus-shifting role. In particular, by imposing an export tax or import subsidy on the final good, the Home government is able to reduce the domestic price of the final-good, while not bearing the full cost of this policy.

With these findings we therefore identify two novel purposes for trade agreements. First, to

[4] Our stylized setup highlights the international nature of the market failure that gives rise to a role for subsidies. Hence, although domestic subsidies in Foreign might also help alleviate the hold-up problem, trade taxes and subsidies will generally continue to be used as part of the set of second-best policies.

the extent that international input trade requires relationship-specific investments between domestic and foreign firms that cannot be perfectly contracted over on an ex-ante basis, the volume of input trade will be too low from the perspective of international efficiency when governments set trade policy unilaterally: by expanding input trade volume a trade agreement can therefore move countries toward the international efficiency frontier. And second, in the presence of such relationship-specific investments in inputs and the bargaining over price that is implied, governments face incentives to distort final-good trade as well, as a means of extracting bargaining surplus from foreign firms: by eliminating these additional policy distortions as well, a trade agreement can help countries achieve the international efficiency frontier. An interesting feature of this second purpose of trade agreements is that it need not imply that trade volumes in final goods should be increased under the agreement: indeed, we identify circumstances under which a country must agree to policy changes which *restrict* trade volume in the final good in an internationally efficient agreement.

This general reasoning stands in sharp contrast to the existing theories of trade agreements, which cast the problem that a trade agreement exists to correct as either stemming from a terms-of-trade driven Prisoners' Dilemma between the governments of countries that are large in world markets (the terms-of-trade theory), or stemming from an inability of governments to make commitments to the private sector (the commitment theory).[5] We make assumptions that preclude a commitment role for trade agreements in this paper. On the other hand, the role that we identify here is related to the terms-of-trade theory, in that there and here the purpose of a trade agreement is to induce those governments that have the ability to affect foreign exporter (international) prices with their trade policy choices to behave in an internationally efficient way. But our theory exhibits some key differences.

First, foreign exporter prices are determined by competitive (or sometimes non-competitive) *market* conditions in the terms-of-trade theory, while here (the relevant) foreign exporter prices are determined as a result of ex-post bilateral producer-supplier *bargaining*. This leads to distinct predictions across the theories concerning the determinants of the degree of trade liberalization required to bring countries to the international efficiency frontier. According to the terms-of-trade theory, it is the Nash foreign export supply elasticities faced by each country that predict the inefficiencies in its Nash tariff choices, and therefore that predict the changes in its policies necessary to reach the international efficiency frontier (see, for example, Bagwell and Staiger, 1999, 2006). In the theory we develop here it is the details of the ex-post bilateral producer-supplier bargaining that predict the inefficiencies in each country's Nash tariff choices: for example, we show that the input-producing Foreign country will have an incentive to use an export tax to appropriate surplus from the Home country only when its suppliers have sufficiently weak bargaining power relative to final-good producers at Home. One implication of these differences is that the terms-of-trade theory predicts that the fundamental task of a trade agreement is always to increase trade volumes from their Nash levels, while as we have observed above it is possible under the theory we develop

[5] See Bagwell and Staiger (2002, Chapter 2) for a description of the major theories of trade agreements.

here that a trade agreement should lead to *reductions* in some (i.e., the final good) trade volumes relative to their Nash levels.

Second, according to the terms-of-trade theory there is a single international problem that a trade agreement must solve, and that is terms-of-trade manipulation, while here there is also a second international problem that a trade agreement must address, namely, the international hold-up problem. According to the terms-of-trade theory, if governments ignored the effect of their policies on world prices, the resulting tariffs would coincide with efficient choices. This feature is not shared by our model, because with relationship-specific investments and incomplete contracting, active manipulation of the terms of trade also serves the valuable role of enhancing the incentives of suppliers to invest. In fact, as long as governments seek to maximize aggregate real national income with their trade policies, the role of an international trade agreement in our model is precisely to nullify the "bad" terms-of-trade manipulation, while *maintaining* the "good" terms-of-trade manipulation inherent in the second-best subsidies to input trade.

A final difference between our theory and the terms-of-trade theory arises when political economy motives are introduced. According to the terms-of-trade theory, preventing governments from manipulating the terms of trade with their trade policy choices is *the* task of a trade agreement, regardless of whether the member governments are concerned only with maximizing aggregate real national income or rather have distributional (e.g., political) concerns as well (see Bagwell and Staiger, 1999). In our model, as we have indicated above, the role of a trade agreement can be interpreted as preventing governments from engaging in "bad" terms-of-trade manipulation provided governments seek to maximize aggregate real national income. But contrary to the terms-of-trade theory, we show that when governments have political economy motives the task of a trade agreement according to the theory we develop here becomes more complicated, and cannot be reduced to solving a simple terms-of-trade problem.

To our knowledge, this is the first paper to consider *endogenous* trade policy choices in a model with intermediate input trade in an environment with relationship-specific investments and incomplete contracting. Ornelas and Turner (forthcoming) develop a model in which import tariffs on intermediate inputs are shown to aggravate the hold-up problem in international vertical relationships, with the implication that trade liberalization may lead to a larger increase in trade flows than in standard models. Ornelas and Turner do not however study optimal trade policies or the possibility of trade agreements in their framework.[6] McLaren (1997) studies the desirability of announcing a future trade liberalization in a model where producers incur sunk costs to service foreign markets, but his framework emphasizes commitment problems from which we completely abstract.[7]

The rest of the paper is organized as follows. In section 2, we develop a Benchmark Model that

[6] Similarly, Antràs and Helpman (2004) and Diez (2006) study the effect of trade frictions on the choice of organizational form of firms contemplating offshoring, but they also treat trade frictions as exogenous.

[7] Yarbrough and Yarbrough (1992) also emphasize commitment problems associated with trade relationships that involve substantial relationship- (or market-) specific investments, but they focus on how these issues affect the choice between unilateral liberalization, bilateral agreements and multilateral agreements.

introduces the international hold-up problem and illustrates the valuable role of active second-best trade policies. In section 3, we consider Nash equilibrium policy choices and show that they leave room for a welfare-enhancing international trade agreement. In section 4, we develop an extension with politically motivated governments and relate our theory of trade agreements to the terms-of-trade theory. In section 5, we consider a variety of extensions of the model. We offer some concluding remarks in section 6.

2 The Benchmark Model

We begin this section by describing a benchmark free-trade two-small-country trade model in which final-good producers in the home country import inputs from suppliers in the foreign country. We refer to this model as the *Benchmark Model*. While simple and special along a number of dimensions, the Benchmark Model is meant to highlight the essential features of the basic international hold-up problem which arises under free trade. After presenting the setup and characterizing the free-trade equilibrium, we derive the (second-best) trade policies that maximize world welfare.

2.1 Setup

We consider a world of two small countries, Home (H) and Foreign (F), and a large rest-of-world whose only role in the model is to fix the prices at which varieties of a final good 1 are available to H and F on world markets (the direction of trade in good 1 is not specified and is immaterial). Consumer preferences are identical in H and F and given by

$$U^j = c_0^j + u\left(c_{1_T}^j + \delta c_{1_G}^j\right), \tag{1}$$

where c_i^j is consumption of good $i \in \{0, 1_T, 1_G\}$ in country $j \in \{H, F\}$, where $u' > 0$ and $u'' < 0$, and where $\delta \in (0, 1)$. Good 0, which we take to be the numeraire, is assumed to be costlessly traded and available in sufficient quantities that it is always consumed in positive amounts in both H and F. Good 1 comes in two types: a customized type T and a generic type G. Note that the preferences in (1) are such that consumers are willing to buy both types of good 1 only if the price of the generic relative to that of the customized type is equal to δ. This is analogous to consumers perceiving the two goods as perfect substitutes up to a quality shifter. By choice of units for measuring the quantity of the customized type of good 1, we set its (fixed) price on world markets equal to 1. For now we assume that trade in good 1 is free, so that the price of the customized type of good 1 is equal to 1 everywhere in the world.

Regardless of the type, good 1 is produced with an intermediate input x according to the production function $y(x)$, with $y'(x) > 0$ and $y''(x) < 0$.[8] The key difference between the two types of good 1 is that the production of a generic good G uses an intermediate input x that is not customized to the producer's needs. We suppose that the home country H is inhabited

[8] In order to ensure that the second-order conditions are met, we will later impose additional assumptions on $y(x)$.

by a unit measure of producers (or entrepreneurs, or distributors) of the final good 1, while the foreign country F is inhabited by a unit measure of suppliers (or managers, or wholesalers) of the intermediate input x. Hence, to produce the final good 1, producers in H must import inputs from suppliers in F. We assume that the marginal cost of input production (generic or customized) in F is constant and, through choice of the units in which inputs are measured, we normalize it to 1. For now, we also assume that trade in x is free.

We next turn to focus on the nature of the bilateral relationship between a final-good producer in H and an input supplier in F, which comprises the essence of the model. We adopt a setting of incomplete contracts between final-good producers and input suppliers. In our Benchmark Model, contractual incompleteness can be rationalized in the following simple way. Following Grossman and Helpman (2002) and Antràs (2003), we assume that, when investing in the supply of (customized or non-customized) x, the supplier can choose between manufacturing a high-quality or a low-quality input, and the latter can be produced at a negligible cost but is useless to final-good producers. The quantity of x is observable to everyone and therefore verifiable by third-parties, but we assume that the quality of x is only observable to the supplier and producer in the particular bilateral relationship, and so quality-contingent contracts are not available. Although parties could still sign a contract specifying a price and a quantity, if they did so, the supplier would always have an incentive to produce the low quality input (at lower cost) and still receive the same contractually stipulated price.[9]

Hence, in this environment, no contracts are signed between suppliers and producers prior to the initial supplier investment decisions. And without an initial contract, the price at which each supplier in F sells its inputs to a producer in H is then decided ex-post (through bargaining) once quality has been chosen. We follow the bulk of the literature in assuming that the bargained price is determined through symmetric Nash bargaining. Because parties have symmetric information at the bargaining stage, ex-post efficiency ensures that low-quality production will never be chosen by an input supplier in equilibrium, and so this dimension of the model can be kept in the background henceforth.

We now describe the structure of the bilateral producer-supplier relationship in detail. We assume that all agents have an ex-ante zero outside option. The sequence of events is as follows:

stage 1. The unit measure of producers in H and suppliers in F are randomly matched, producing a unit measure of matches.[10] Each agent decides whether to stay with their match or exit the market. In the former case, the producer provides the supplier with a list of customized input specifications. In the latter case, each producer obtains their ex-ante outside option (equal to zero).

stage 2. Each supplier decides on the amount x of customized input to be produced (at marginal

[9] There is a large literature proposing a variety of mechanism-design resolutions to the hold-up inefficiencies caused by incomplete contracts. These resolutions however generally rely on the ability of parties to commit not to renegotiate an initial contract. Bolton and Dewatripont (2005, Chapter 12) offer an excellent overview of the insights and limitations of this literature.

[10] We introduce search frictions later in the paper.

cost of 1) – the marginal cost of production for a generic input (not customized to the matched producer's needs) is the same, so there is no benefit in not customizing the input for the matched producer at this point.

stage 3. Each producer-supplier pair bargains over the price of the intermediate input (we assume symmetric Nash bargaining).

stage 4. A small number (formally, a measure-zero countable infinity) n of the bilateral pairs are exogenously dissolved and randomly rematched in a secondary market. They bargain again, with each party capturing again one-half of the gains from trade. No further inputs can be produced; the amount produced in *stage* 2 is perceived as generic in the secondary market because it was tailored to another producer's specifications with probability one.

stage 5. Each producer in H imports x from its partner-supplier and produces the final good with the acquired x, and payments agreed in *stages* 3 and 4 are settled.

This 5-stage game generates the simple hold-up problem that forms the heart of our analysis. A number of features of this setup are worth noting at this point.

First, we rule out the use of ex-ante (*stage*-1) lump-sum transfers between producers and suppliers. The possibility of these transfers is particularly hard to defend in the international context that we study, where such transfers and the obligations associated with them might be difficult to enforce. In section 5, however, we will show that our main results are robust to allowing for these transfers.

Second, the role of *stage* 4 is to pin down the outside options of the producer and the supplier should their *stage*-3 bargaining break down. In light of the broader structure of our benchmark free-trade model, it is easy to see that the breakup of a single bargaining pair in *stage* 3 would result in each member of the pair being rematched with probability 1 with a random partner in *stage* 4, and therefore that *stage* 4 implies an outside option equal to $\frac{1}{2}\delta y(x)$ for both the producer in H and the supplier in F as they engage in Nash bargaining in *stage* 3. Beyond this, *stage* 4 plays no role, and in particular only the customized type of good 1 will be produced with positive measure in equilibrium.

Finally, we note that production-side efficiency requires that the intermediate input is used to produce the customized final good T, and that the customized input is produced at a level x^E which satisfies

$$y'\left(x^E\right) = 1, \tag{2}$$

and thereby equates the marginal revenue generated from an additional unit of the input (recall that the price of the customized final good is fixed by world markets and equal to 1 under free trade) with the marginal cost of producing an additional unit of the input (which is constant and normalized to 1).

2.2 Free Trade Equilibrium

We now characterize the subgame perfect equilibrium of the 5-stage game described in the previous subsection. The characterization follows very simply from a few key observations. We consider a representative producer in H and supplier in F that are matched in *stage* 1.

First, if the producer uses the supplier's input to produce the final good in *stage* 5, its revenue is given by $y(x)$. Second, as observed in the previous subsection, the outside options of both the producer and the supplier in their *stage*-3 Nash bargain are $\frac{1}{2}\delta y(x)$, so the sum of their outside options is equal to $\delta y(x)$. Hence, the quasi-rents over which the producer and supplier bargain in *stage* 3 (recall that the cost of producing x is sunk at this point) are $(1-\delta)\, y(x)$. Therefore, in the symmetric Nash bargain of *stage* 3, the final-good producer in H and the input supplier in F both obtain a payoff of $\frac{1}{2}y(x)$.

Next, rolling back to *stage* 2, observe that the input supplier chooses x to maximize $\frac{1}{2}y(x) - x$, so the optimal quantity $\hat{x}$ of input satisfies

$$\frac{1}{2}y'(\hat{x}) = 1. \tag{3}$$

Given the concavity of $y(x)$, it is clear from a comparison of (3) with (2) that $\hat{x} < x^E$. This is the under-investment associated with the hold-up problem, and it reflects the fact that the producer and supplier bargain over the price of the input *after* the supplier has already sunk investment in input supply.[11]

Finally, consider *stage* 1. If the producer hands the supplier a list of customized input specifications, the producer anticipates obtaining a payoff equal to

$$\pi^H = \frac{1}{2}y(\hat{x}),$$

which exceeds the payoff it would obtain by not providing the specifications.[12] Similarly, by accepting to form a partnership with the home producer, the supplier anticipates obtaining a payoff of

$$\pi^F = \frac{1}{2}y(\hat{x}) - \hat{x},$$

which also exceeds its ex-ante outside option.[13] In sum, no separations will occur at stage 1. Note also that the sum of payoffs of the two parties is equal to $y(\hat{x}) - \hat{x}$, which is strictly less than the sum of payoffs that would obtain when investment is chosen at the efficient level x^E defined by (2).

[11] The fact that the size of the inefficiency is independent of δ follows from our assumption of symmetric Nash bargaining in the secondary market. See section 5.1 for the case with generalized Nash bargaining.

[12] Remember that the ex-ante outside option of producers is equal to 0. We could allow producers that choose to exit at stage 0 to later access the secondary market at *stage* 4. But note that, at most, these producers would obtain a payoff equal to $\delta\pi^H < \pi^H$ in that secondary market, so again it is optimal for the producer to stay in the relationship.

[13] Given the concavity of $y(x)$, we have $\frac{1}{2}y(\hat{x}) - \hat{x} \geq \frac{1}{2}xy'(\hat{x}) - \hat{x} = 0$. Furthermore, we also have $\frac{1}{2}y(\hat{x}) - \hat{x} \geq \frac{1}{2}\delta y(\hat{x}) - \hat{x}$, where the right-hand-side is the suppliers' payoff when allowing them to enter the secondary market at *stage* 4 and finding a match with probability one.

Now consider the measure of social welfare in each country implied by our Benchmark Model. With our assumption of quasilinear preferences, this measure is given by consumer surplus plus profits plus trade tax revenue (the latter being zero under free trade).[14] Using (1), we have that country j's demand for good 1 is given by $D_1(p_1^j) \equiv u'^{-1}(p_1^j)$, with consumer surplus then defined as $CS^j(p_1^j) \equiv \int_{p_1^j}^{\bar{p}} D_1(p)dp$ where $\bar{p}$ is the "choke" price for country j's demand of good 1. World aggregate welfare may then be represented by

$$W^W = W^H + W^F = CS^H(1) + CS^F(1) + \pi^H + \pi^F = CS^H(1) + CS^F(1) + y(\hat{x}) - \hat{x},$$

which is strictly lower than world welfare in the presence of production efficiency because $y(\hat{x}) - \hat{x} < y\left(x^E\right) - x^E$. We summarize this discussion with:

Proposition 1 *In the Benchmark Model, a hold-up problem exists under free trade, leading to insufficient investment in the production of imported foreign inputs ($\hat{x} < x^E$). As a result, aggregate world welfare is inefficiently low.*

We next turn to consider trade intervention as a possible means of alleviating the hold-up problem.

2.3 Second-Best Trade Policy

We now explore the possible beneficial role of trade policy in this distorted economy. To this end, we let τ_x^H denote the trade tax imposed by H on imports of the input x (positive if an import tariff, negative if an import subsidy) defined in specific terms, and we let τ_x^F be the analogous trade tax imposed by F. Furthermore, we let τ_1^H denote the trade tax imposed by H on the home country's trade in the final good 1 (positive if an import tariff or export subsidy, negative if an import subsidy or export tax) also defined in specific terms. Observe that the price of the final good 1 in H is now given by $p_1^H = 1 + \tau_1^H$, whereas the price of the input x continues to be determined by Nash bargaining between producers and suppliers (though trade taxes may affect this negotiated price).[15]

How does the introduction of these trade taxes affect the equilibrium characterized in the previous subsection? To explore this question, we first consider the case of second-best trade policies, that is, the set of policies that maximize aggregate world welfare (subject to the contractual frictions in producer-supplier relationships). More specifically, we introduce the following *stage* 0 which occurs prior to *stage* 1 of the 5-stage game described in subsection 2.1:

[14] Strictly speaking, social welfare should also include a term related to income earned by other factors of production (say labor) in the economy. Nevertheless, it would be straightforward to close the model in a way that makes this term independent of policies in sector 1 (see, for instance, Grossman and Helpman, 1994). Henceforth, we simply ignore this term.

[15] We could have also allowed for a final-good trade tax τ_1^F imposed by F, but it is intuitively clear (and is easily shown) that there will be no incentive to use such an instrument, since such trade taxes could only alter the domestic price of good 1 in F (owing to F's small size on world markets) and that price has no impact on the hold-up problem between F's input suppliers and H's final good producers.

stage 0. A social planner selects a home-country trade tax τ_1^H on the final good 1, a home-country import tax τ_x^H on home imports of the input x, and a foreign-country export tax τ_x^F on foreign exports of the input x.

After the social planner has selected these import tariffs/subsidies in *stage* 0, the sequence of events is as outlined in subsection 2.1 (with trade taxes collected at the time of importation and production/sales in *stage* 5).

Consider now how these trade policy choices in *stage* 0 affect the equilibrium outcome of the game. In their *stage*-3 bargaining, if the producer and supplier reach an agreement they stand to obtain a joint payoff of (recalling again that the cost of producing x is sunk at that point)

$$\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x.$$

A positive import tariff or export subsidy on the final good ($\tau_1^H > 0$) raises the joint surplus of the producer and supplier because it raises the price at which the final good is sold in H. Conversely, a positive import tariff ($\tau_x^H > 0$) or export tax ($\tau_x^F > 0$) on inputs reduces the joint surplus of the producer and supplier because it transfers part of the surplus to governments.

If the producer and the supplier do not reach an agreement, the final-good producer and its input supplier are each left with a *stage*-4 payoff equal to

$$\frac{1}{2}\left(\delta\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x\right).$$

These expressions are valid provided they are non-negative, and here and throughout the body of the paper we characterize results for the case where these non-negativity constraints are non-binding. In the Appendix we consider the cases where a non-negativity constraint is binding (so that the associated payoff is zero), and show that our qualitative results carry through for those cases as well.

Notice that the ex-post gains from agreement (or quasi-rents) are now given by $(1-\delta)\left(1+\tau_1^H\right) y\left(x\right)$. Hence, both parties obtain a payoff equal to $\frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x\right)$ in the Nash bargain of *stage* 3, and the input supplier's choice of x in *stage* 2 must then satisfy

$$\frac{1}{2}\left(1+\tau_1^H\right) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\left(\tau_x^H + \tau_x^F\right). \quad (4)$$

It is clear from (4) that $\hat{x}$ is increasing in τ_1^H and decreasing in τ_x^H and τ_x^F. Intuitively, incomplete contracting leads to rent-sharing between the producer and supplier, and hence the latter's incentives to invest tend to be higher whenever the surplus is higher, that is when τ_1^H is higher and when τ_x^H or τ_x^F are lower. We will see in later sections that the positive dependence of $\hat{x}$ on τ_1^H and negative dependence of $\hat{x}$ on τ_x^H and τ_x^F hold for a variety of specifications of the game played between the producer and supplier.

At *stage* 1, the final-good producer in H anticipates a payoff equal to

$$\pi^H = \frac{1}{2}\left(\left(1+\tau_1^H\right)y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right)\hat{x}\right), \tag{5}$$

while the supplier in F expects a payoff equal to

$$\pi^F = \frac{1}{2}\left(\left(1+\tau_1^H\right)y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right)\hat{x}\right) - \hat{x}, \tag{6}$$

where $\hat{x}$ is implictly defined by (4). As a result, welfare at home, inclusive of tax revenue, is given by

$$W^H = CS^H(1+\tau_1^H) + \pi^H + \tau_1^H\left[D_1(p_1^H) - y\left(\hat{x}\right)\right] + \tau_x^H\hat{x},$$

while welfare in foreign is

$$W^F = CS^F(1) + \pi^F + \tau_x^F\hat{x}.$$

We now seek to characterize the set of trade policy choices that maximize world welfare. Formally, we are seeking the triplet $\left(\tau_1^{HE}, \tau_x^{HE}, \tau_x^{FE}\right)$ that maximizes (see eq. (5) and (6)):

$$W^W = W^H + W^F = CS^H(1+\tau_1^H) + CS^F(1) + \left(1+\tau_1^H\right)y\left(\hat{x}\right) - \hat{x} + \tau_1^H\left[D_1(1+\tau_1^H) - y\left(\hat{x}\right)\right],$$

subject to $\hat{x}$ being given by (4).[16] Notice that for a given value of the sum $\tau_x^H + \tau_x^F$, the individual values of τ_x^H and τ_x^F have no effect on world welfare. This implies that the second-best policies will only pin down an aggregate input trade tax $\tau_x \equiv \tau_x^H + \tau_x^F$. The efficient policies τ_1^{HE} and τ_x^E are then determined by the following first-order conditions of the problem above:[17]

$$\begin{aligned} \frac{\partial W^W}{\partial \tau_1^H} &= \tau_1^H \frac{\partial D_1}{\partial p_1^H} + \left[y'\left(\hat{x}\right) - 1\right]\frac{\partial \hat{x}}{\partial \tau_1^H} = 0, \text{ and} \\ \frac{\partial W^W}{\partial \tau_x} &= \left[y'\left(\hat{x}\right) - 1\right]\frac{\partial \hat{x}}{\partial \tau_x} = 0. \end{aligned} \tag{7}$$

The first-order conditions in (7) are instructive. Recalling from (4) that $\partial\hat{x}/\partial\tau_1^H > 0$, it is clear from (7) that the optimal choice of τ_1^H is strictly positive, provided that $\left[y'\left(\hat{x}\right) - 1\right] > 0$ which by (2) implies that $\hat{x} < x^E$: this suggests that an import tariff or export subsidy on trade in the final good 1 could raise welfare in H, by increasing $\hat{x}$ toward x^E and thereby helping to ameliorate the hold-up problem at the cost of lost consumer surplus. However, recalling from (4) that $\partial\hat{x}/\partial\tau_x^H < 0$, it is clear from (7) that the optimal choice of τ_x^H must ensure that $\left[y'\left(\hat{x}\right) - 1\right] = 0$, thereby achieving productive efficiency: there is no associated loss in consumer surplus when the tariff on imported inputs τ_x^H is used to increase $\hat{x}$, and the optimal choice of τ_x^H therefore solves completely the hold-up problem and achieves productive efficiency. This in turn leaves no reason for government intervention with regard to trade in the final good 1. Hence, the optimal choice

[16] It is the presence of this constraint that leads us to refer to $\left(\tau_1^{HE}, \tau_x^{HE}, \tau_x^{FE}\right)$ as *second-best* trade policy choices, although we shall see that these policy choices will lead to an attainment of the first-best welfare level.

[17] It is easily checked that second-order conditions are satisfied (see Appendix).

of τ_1^H, which we denote by τ_1^{HE}, is $\tau_1^{HE} = 0$. On the other hand, the second-best policies do call for intervention with regards to input trade. In particular, from equation (4) it follows that the optimal trade tax is an input subsidy in an amount equal to $\tau_x^E \equiv \tau_x^{HE} + \tau_x^{FE} = -1$. We may thus state:

Proposition 2 *In the Benchmark Model, the second-best trade policy choices maintain free trade in the final good and subsidize importation of the input so as to solve the hold-up problem and achieve productive efficiency.*

The intuition for Proposition 2 is simple. The hold-up problem between producers in H and suppliers in F results in a level of imported inputs which is inefficiently low. The market failure is an international one in nature, and thus it is natural that trade taxes or subsidies can serve a useful role in alleviating the inefficiency. Furthermore, although trade intervention in the final good could be used to raise the home-country price of the final good and increase the volume of imported inputs (through rent-sharing), this would come at a cost of reduced home-country consumer surplus. A subsidy to imported inputs does not reduce consumer surplus, but it nevertheless succeeds in increasing the volume of imported inputs by increasing the surplus over which the parties negotiate in the ex-post (*stage*-3) bargain. As a consequence, a subsidy to imported inputs targets just the distorted margin, and in analogy with the targeting principle (Bhagwati and Ramaswami, 1963, Johnson, 1965) is hence the first-best method of addressing the problem.

We have thus identified a novel role for trade policy intervention, namely, as a means of addressing the international hold-up problem that arises when international trade requires relationship-specific investments between domestic producers and their foreign suppliers. A natural question is whether the *unilateral* trade policy choices of both the home and foreign governments will lead to overall trade distortions that concord with the efficiency conditions outlined in Proposition 2. We tackle this issue in the next section.

3 Nash Equilibrium Trade Policy Choices

In this section we characterize the Nash policies of the home and foreign governments and evaluate the potential role of trade agreements in our Benchmark Model. In order to build intuition, we first consider the unilaterally optimal trade policy choices of the home government when the foreign government follows a policy of free trade, and only later consider the possibility of foreign trade policy intervention.

3.1 Unilateral Home Policy

We first consider the subgame perfect equilibrium of the Benchmark Model for the case in which *stage* 0 is as follows:

stage 0. The home government H selects a trade tax τ_1^H on the final good 1, and a trade tax τ_x^H on the imported input x; the foreign government F remains passive, i.e., $\tau_x^F = 0$.

Following the same steps as in the last section, and setting $\tau_x^F = 0$ at *stage* 0, we have that the final-good producer in H now obtains a *stage*-2 payoff equal to:

$$\pi^H = \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H x\right), \tag{8}$$

where $\hat{x}$ is now implictly defined by

$$\frac{1}{2}\left(1+\tau_1^H\right) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\tau_x^H. \tag{9}$$

With these expressions in hand, home welfare can be written as the sum of home consumer surplus, profits and tax revenue, or

$$W^H = CS(1+\tau_1^H) + \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H \hat{x}\right) + \tau_1^H [D_1(1+\tau_1^H) - y\left(\hat{x}\right)] + \tau_x^H \hat{x}.$$

The optimal choice of τ_1^H and τ_x^H, which we denote by $\hat{\tau}_1^H$ and $\hat{\tau}_x^H$, must maximize home welfare W^H, and will hence satisfy the first-order conditions

$$\begin{aligned}
\frac{\partial W^H}{\partial \tau_1^H} &= 0 = \tau_1^H \frac{\partial D_1}{\partial p_1^H} - \frac{1}{2} y\left(\hat{x}\right) + \left[\frac{1}{2}\left(1-\tau_1^H\right) y'\left(\hat{x}\right) + \frac{1}{2}\tau_x^H\right] \frac{\partial \hat{x}}{\partial \tau_1^H}, \text{ and} \\
\frac{\partial W^H}{\partial \tau_x^H} &= 0 = \frac{1}{2}\hat{x} + \left[\frac{1}{2}\left(1-\tau_1^H\right) y'\left(\hat{x}\right) + \frac{1}{2}\tau_x^H\right] \frac{\partial \hat{x}}{\partial \tau_x^H},
\end{aligned}$$

where recall that $\hat{x}$ is given by equation (9).[18] Applying the implicit function theorem (twice) to (9) delivers

$$\frac{\partial \hat{x}/\partial \tau_1^H}{\partial \hat{x}/\partial \tau_x^H} = -y'\left(\hat{x}\right),$$

which can be used to manipulate the above first-order conditions to obtain:

$$\begin{aligned}
\hat{\tau}_1^H &= \frac{\frac{1}{2}\hat{x}[y'\left(\hat{x}\right) - \frac{y(\hat{x})}{\hat{x}}]}{-\partial D_1/\partial p_1^H}, \text{ and} \\
\hat{\tau}_x^H &= -(1-\hat{\tau}_1^H) y'\left(\hat{x}\right) - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H}.
\end{aligned} \tag{10}$$

The expressions in (10) reflect an interesting logic. As in the case of second-best policies, part of the goal of the home government in intervening with τ_1^H and/or τ_x^H is to raise $\hat{x}$ towards its efficient level x^E. Nevertheless, the home government *does not* maximize world welfare and hence there is an offsetting leakage of surplus to the foreign supplier that must be taken into account by the home government in setting its optimal unilateral policies. This leads to two observations: first, it is no longer optimal to deliver the chosen $\hat{x}$ using only τ_x^H, and the use of τ_1^H reflects a new and independent source of international inefficiency associated with the unilateral policy choices of

[18] The second-order conditions for this problem do not reduce to simple expressions, as was the case with second-best policies. In the Appendix, we discuss these second-order conditions and show that they are satisfied for a simple parameterized example.

the home country; and second, it is no longer optimal for the home country to raise $\hat{x}$ all the way to its efficient level x^E.

The first observation can be understood as follows. The home government must now concern itself with two tasks as it considers its policy choices. First, it must face foreign suppliers with the appropriate *marginal* incentives for investment in the supply of x so as to achieve the desired investment level $\hat{x}$. Second, the home government must also concern itself with extracting *infra-marginal* surplus from foreign suppliers through the use of trade policy instruments. With its two tariff instruments τ_1^H and τ_x^H, the home government can orchestrate infra-marginal foreign surplus extraction with adjustments in τ_1^H and τ_x^H that hold $\hat{x}$ fixed, so that $d\tau_x^H(\tau_1^H)/d\tau_1^H = -\frac{\partial\hat{x}/\partial\tau_1^H}{\partial\hat{x}/\partial\tau_x^H}$, and can extract foreign surplus in this fashion at the rate

$$\frac{\partial W^F(\tau_1^H, \tau_x^H(\tau_1^H))}{\partial\tau_1^H}|_{d\hat{x}=0} = -\frac{1}{2}\hat{x}\left[y'(\hat{x}) - \frac{y(\hat{x})}{\hat{x}}\right]. \tag{11}$$

Evidently, with the concavity of $y(x)$ implying $[y'(\hat{x}) - \frac{y(\hat{x})}{\hat{x}}] < 0$, it follows from (11) that for any given level of $\hat{x}$, additional surplus can be extracted from the foreign country by reducing τ_1^H and accompanying this with a reduction in τ_x^H which preserves the level of $\hat{x}$. What, then, prevents the home country from lowering τ_1^H in this fashion indefinitely, until all of the surplus has been extracted from foreign suppliers? The impact on home-country welfare of these tariff changes is given by

$$\frac{\partial W^H(\tau_1^H, \tau_x^H(\tau_1^H))}{\partial\tau_1^H}|_{d\hat{x}=0} = \tau_1^H\frac{\partial D_1^H}{\partial p_1^H} + \frac{1}{2}\hat{x}\left[y'(\hat{x}) - \frac{y(\hat{x})}{\hat{x}}\right]. \tag{12}$$

As (12) makes clear, what eventually stops this process of foreign surplus extraction is the growing home-country final-good demand distortions that are associated with $\tau_1^H < 0$. It is for these reasons that (10) implies $\hat{\tau}_1^H < 0$: in words, it is unilaterally optimal for the home government to utilize trade policy to distort downward the price of the final good 1 in the home market (through either an import subsidy or an export tax on the final good) as a means of extracting bargaining surplus from foreign suppliers. Notice as well that this incentive is present regardless of the level of $\hat{x}$, and in particular is present even when $\hat{x}$ is set at its efficient level x^E. Hence, our model identifies a new and independent source of international inefficiency when the home country sets its tariffs unilaterally: the attempt to extract bargaining surplus from foreign suppliers by distorting the home market price of the final good 1.

The second observation above, that it is no longer optimal for the home country to raise $\hat{x}$ all the way to its efficient level x^E, can be confirmed by considering the expression for $\hat{\tau}_x^H$ in (10). This expression is of indeterminate sign, indicating that $\hat{\tau}_x^H$ can now be either negative (an import subsidy on inputs of x) or positive (an import tariff on inputs of x): this reflects the tension that arises for the home-country government between correcting the hold-up problem and capturing surplus from the foreign input supplier, a tension that was absent in the choice of second-best policies in section 2.3. To see this formally, substitute (4) into the expression for $\hat{\tau}_x^H$ in (10) and simplify to

$$y'(\hat{x}) = 1 - \frac{1}{2}\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H} > 1, \tag{13}$$

which implies that $\hat{x} < x^E$. Hence, at least when the foreign government remains passive, it is unilaterally optimal for the home government to utilize its trade policies in a way that does *not* fully correct the international hold-up problem.

We can thus conclude that when only H intervenes international efficiency is not achieved. Instead, there are now *two* sources of international inefficiency that arise: an inefficiently low input trade volume that results from the continued existence of the international hold-up problem; and distortions in the final good market that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers. We may thus state:

Proposition 3 *In the Benchmark Model, when only H intervenes with trade policy, its unilaterally optimal policy choices reduce the local price of the final good – with an import subsidy or an export tax on trade in good 1 – and may either tax or subsidize importation of the input, but in any case result in (i) insufficient investment in the production of imported foreign inputs ($\hat{x} < x^E$) and the continued existence of the international hold-up problem, and (ii) distortions in the final good market that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers.*

Proposition 3 stands in marked contrast to Proposition 2, and reflects a simple point. To the extent that home-country producers share part of the surplus from production with foreign suppliers (as is the case in our Benchmark Model), the unilateral incentives of the home-country government to intervene with trade policy to mitigate the international hold-up problem will be muted by the fact that foreign suppliers enjoy some of the benefits of this intervention. In this environment, the home-country's unilateral intervention must be concerned as well with capturing foreign surplus, and therefore the home country cannot be counted on to solve the international hold-up problem on its own. Moreover, the home-country's attempts to extract bargaining surplus from foreign suppliers will spill-over into the final good market as well, and introduce additional distortions there.

3.2 Foreign Intervention and Nash Equilibrium Policy Choices

We turn next to consider the unilateral incentives of the foreign government to intervene with a trade tax τ_x^F (as before, in a prior *stage* 0). We hence modify stage 0 as follows:

stage 0. The home government H selects a trade tax τ_1^H on the final good 1, and a trade tax τ_x^H on the imported input x; simultaneously, the foreign government F selects a trade tax τ_x^F on the exported input x.[19]

[19] Implied by the timing of the governments' tariff choices is the assumption that the governments can make tariff commitments to the private sector. If the governments did not have this ability, then as is well known a separate

As in the case of second-best policies, we could allow for foreign taxes on trade in the final good 1, but these have no effect on the hold-up problem and will thus never be used as a part of an optimal set of policies. But what about an export tax/subsidy imposed by F on its exports of inputs to H? Presumably, an export subsidy on F's exports of x to H could help solve the hold-up problem, but it is unclear that F would have an incentive to intervene in this way when H's producers capture part of the surplus from the bilateral producer-supplier relationship. Furthermore, it is not obvious that F's government will have an incentive to use trade policy to manipulate the choice of x, when this choice is made by foreign suppliers to maximize the profits that accrue to foreigner residents. A more promising possibility is that F might benefit by taxing its exports of x, if in doing so it can collect revenue from producers in H, while not bearing the full cost of the induced distortions.

To sort out these possibilities, we consider F's incentive to intervene facing a given H policy pair $\left(\tau_1^H, \tau_x^H\right)$. In this case, the input supplier in F now has a payoff of

$$\pi^F = \frac{1}{2}(1+\tau_1^H)y(\hat{x}) - (1 + \frac{1}{2}\tau_x^H + \frac{1}{2}\tau_x^F)\hat{x},$$

with $\hat{x}$ defined by

$$\frac{1}{2}(1+\tau_1^H)y'(\hat{x}) = 1 + \frac{1}{2}\tau_x^H + \frac{1}{2}\tau_x^F. \tag{14}$$

Foreign welfare is then given by the sum of foreign consumer surplus, profits and tax revenue:

$$W^F = CS^F(1) + \frac{1}{2}\left(1+\tau_1^H\right)y(\hat{x}) - (1 + \frac{1}{2}\tau_x^H + \frac{1}{2}\tau_x^F)\hat{x} + \tau_x^F\hat{x}.$$

The optimal choice of τ_x^F, which we denote by $\hat{\tau}_x^F$, hence must satisfy the first-order condition

$$\frac{\partial W^F}{\partial \tau_x^F} = 0 = \frac{1}{2}\hat{x} + [\frac{1}{2}\left(1+\tau_1^H\right)y'(\hat{x}) - 1 - \frac{1}{2}\tau_x^H + \frac{1}{2}\tau_x^F)]\frac{\partial \hat{x}}{\partial \tau_x^F}. \tag{15}$$

Recalling that $\partial\hat{x}/\partial\tau_x^F < 0$, the first-order condition in (15) together with (14) immediately implies that

$$\tau_x^F = -\frac{1}{2}\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F} > 0, \tag{16}$$

and hence, the Foreign country finds it optimal to set an export tax on the intermediate input.

The logic behind this result was outlined above. First, while a negative foreign export tax could be used to increase the level of x chosen by the foreign supplier, this level is already chosen by the foreign supplier to maximize foreign profits, and so there is no gain to the foreign country from manipulating this choice with export-sector intervention. Second, and as long as foreign suppliers have less than full bargaining power in the secondary market, they will not bear the full cost of

commitment role for trade agreements might arise (see Bagwell and Staiger, 2002, Chapter 2, for a review of this literature). The particular commitment problems that governments face when trade requires relationship-specific investments are emphasized by Yarbrough and Yarbrough (1992) as providing a reason for trade agreements to exist, and by McLaren (1997) as creating the possibility of perverse negotiating outcomes. Our assumed timing permits us to abstract from the possible commitment role of trade agreements throughout this paper, so that we may focus on other issues.

the increase in the marginal cost of production associated with an export tax. In other words, the foreign government will be able to pass part of the cost of the export tax on to the home country while keeping the entire benefit from it (in the form of tax revenue). As a result, the optimal export tax is positive.

How will the home country respond to the setting of an export tax by Foreign? In order to derive the Nash Equilibrium policy choices of the Home country, we next solve for the final-good tax τ_1^N and the input tax τ_x^N that maximize home welfare for a given Foreign policy choice τ_x^F. This pair $\left(\hat{\tau}_1^{HN}, \hat{\tau}_x^{HN}\right)$ will thus maximize

$$W^H = CS(p_1^H) + \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H \hat{x} - \tau_x^F \hat{x}\right) + \tau_1^H [D_1(p_1^H) - y\left(\hat{x}\right)] + \tau_x^H \hat{x}$$

subject to $\hat{x}$ being given by (9). Manipulating the first-order conditions delivers the following two conditions analogous to those in (10):

$$\begin{aligned} \hat{\tau}_1^{HN} &= \frac{\frac{1}{2}\hat{x}\left[y'\left(\hat{x}\right) - \frac{y(\hat{x})}{\hat{x}}\right]}{-\partial D_1/\partial p_1^H}, \text{ and} \\ \hat{\tau}_x^{HN} &= -(1-\hat{\tau}_1^{HN})y'\left(\hat{x}\right) - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H} + \hat{\tau}_x^{FN}, \end{aligned}$$

where we have replaced τ_x^F with $\hat{\tau}_x^{FN} \equiv -\left(1/2\right)\hat{x}/\left(\partial \hat{x}/\partial \tau_x^F\right)$ reflecting the fact that, in the Nash equilibrium, the Foreign government will best respond to the policy choices of Home.

The first equation implies that $\hat{\tau}_1^{HN}$ is again negative, while the second indicates that the sign of $\hat{\tau}_x^{HN}$ is indeterminate. This parallels the results we obtain in the case without foreign retaliation and the intuition is the same as that oulined above. The only difference is the additional term $\hat{\tau}_x^{FN} > 0$ in the second equation, which suggests that overall input taxes will be higher in the case with foreign retaliation, thus aggravating the hold-up problem. This can be confirmed by combining these expressions with equations (14) and (16) to obtain:

$$y'\left(\hat{x}\right) = 1 - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H}.$$

It is then clear that the Nash equilibrium involves suboptimal trade in intermediate inputs, $\hat{x} < x^E$. Furthermore, comparing this to the distortion in the case without foreign retaliation (see eq. (13)), we observe that the gap between $y'\left(\hat{x}\right)$ and 1 will be higher in the case with foreign retaliation.[20] In sum, we have shown that:

Proposition 4 *In the Nash equilibrium of the Benchmark Model, F maintains free trade in the final good and taxes the exports of the input, while H adopts trade intervention in both the final good and the intermediate input. There is insufficient investment in the production of imported foreign inputs ($\hat{x} < x^E$) that results from the continued existence of the international hold-up problem,*

[20] Strictly speaking, we need to appeal to the second-order conditions of the problem to formally prove this result. See the Appendix for details.

and there are distortions in the home-country final good market that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers.

It is worth emphasizing that the home government chooses to distort the price of the final good in its local market, reflecting an underlying incentive that: (i) operates even though the home country is small in world markets for the final good; and (ii) operates even if the home government makes other policy adjustments to hold its input trade volume fixed. What accounts for this? After all, (i) implies directly that there can be no terms-of-trade gains for the home country in the final good market, and under market clearing (ii) would imply that movements in the foreign exporter price of traded inputs are ruled out and therefore could not be a source of terms-of-trade gains for the home country either. The answer is that the foreign exporter price of the input is determined by *bargaining*, not market clearing, and this severs the link between input price and input trade volume that market clearing would have implied, allowing the foreign exporter price of the input to be altered by changes in home-country policies (and thus allowing bargaining surplus to be extracted from the foreign supplier) even when those changes leave input trade volume fixed.

3.3 The Role of a Trade Agreement

As Proposition 4 records, our Benchmark Model indicates that international hold-up problems between home producers and foreign suppliers can give rise to inefficiencies that are not fully corrected by government intervention in non-cooperative settings, hence pointing to a reason for mutually beneficial trade agreements. In effect, as we have demonstrated, each government has the ability to correct the international hold-up problem with its unilateral trade policy choices, but neither government acting alone can capture fully the benefits from solving the problem, and so neither government has sufficient incentive in the Nash equilibrium to fully correct the international hold-up problem and bring the world to the international efficiency frontier.

In our Benchmark Model, then, a reason for a trade agreement between H and F exists, and its purpose is two-fold: to better help governments solve the international hold-up problem between producers and suppliers and thereby expand input trade volume to internationally efficient levels; and to prevent the home government from distorting its final-good market in an effort to extract bargaining surplus from foreign firms. The design of the trade agreement is relatively simple: it should impose free trade in the final good and should constraint the sum of input trade taxes to not exceed -1, that is $\tau_x^H + \tau_x^F \leq -1$.[21] In sum,

Proposition 5 *In our Benchmark Model, a trade agreement that imposes free trade in the final good ($\tau_1^H = 0$) and specifies input taxes such that $\tau_x^H + \tau_x^F = -1$ increases world welfare and brings H and F to the international efficiency frontier. This is achieved by (i) solving the international hold-up problem, and (ii) preventing governments from introducing distortions in an effort to extract*

[21] Note in this setting that the particular values of τ_x^H and τ_x^F can always be chosen to ensure that both countries are better off under the agreement than under the status quo with Nash tariffs, which implies that a successful agreement does not require the exchange lump sum transfers.

bargaining surplus from foreign firms. Furthermore, the agreement does not require the exchange of lump-sum transfers.

As characterized in Proposition 5, the role of a trade agreement in the Benchmark model stands in sharp contrast to prevailing theories, which cast the problem that a trade agreement exists to correct as either stemming from a terms-of-trade driven Prisoners' Dilemma between the governments of countries that are large in world markets (the terms-of-trade theory), or stemming from an inability of governments to make commitments to the private sector (the commitment theory).[22] As previously mentioned (see note 19), we have made assumptions that preclude a commitment role for trade agreements in this paper. The role that we identify here is related to the terms-of-trade theory, in that there and here the purpose of a trade agreement is to induce those governments that have the ability to affect foreign exporter (international) prices with their trade policy choices to behave in an internationally efficient way, but there are also some key differences: for one, foreign exporter prices are determined by competitive (or sometimes non-competitive) market conditions in the terms-of-trade theory, while here, as we have emphasized, (the relevant) foreign exporter prices are determined as a result of ex-post bilateral producer-supplier bargaining; for another, according to the terms-of-trade theory there is a single international problem that a trade agreement must solve, and that is terms-of-trade manipulation, while here there is also a second international problem that a trade agreement must address, namely, the international hold-up problem.

These differences are substantive, and they lead to distinct predictions across the theories concerning the determinants of the degree of trade liberalization required to bring countries to the international efficiency frontier: according to the terms-of-trade theory, the Nash foreign export supply elasticities faced by each country are crucial in determining the inefficiencies in its Nash tariff choices and the implied changes in policies necessary to reach the international efficiency frontier (see, for example, Bagwell and Staiger, 1999, 2006). Instead, in the theory we develop here it is the parameters that govern the ex-post bilateral producer-supplier bargaining that predict the inefficiencies in each country's Nash tariff choices. As one stark example of this difference, the terms-of-trade theory predicts that the fundamental task of a trade agreement is always to increase trade volumes from their Nash levels, while according to Proposition 5 the home country must agree to policy changes which increase the price of the final good 1 in its local markets, and this would amount to an agreement to *restrict* trade volume in the final good 1 if the home country is an importer of this good.[23]

As we will next demonstrate, by undertaking a more systematic evaluation of the differences between the terms-of-trade theory and the theory we present here, it is possible to more formally

[22] See Bagwell and Staiger (2002, Chapter 2) for a description of the major theories of trade agreements.

[23] In the Appendix, we show that when the location of the secondary market is instead in the foreign country, the foreign government now also has an incentive to manipulate the local price of final good 1 in its market as a means of shifting surplus toward its input supplier firms, but seeks to raise rather than lower this price with an import tariff or export subsidy. In this case, if the foreign country is an exporter of the final good 1, a trade agreement that achieves the efficiency frontier would restrict the foreign country's trade volume in that good by prohibiting export subsidies.

identify several novel features of the role of a trade agreement in our model. This evaluation, though, is best carried out once the possibility of politically-motivated governments is entertained. In the next section we extend the Benchmark Model to incorporate these motives, and then return to a comparison between our theory and the terms-of-trade theory in order to provide a systematic evaluation of their differences.

4 The Benchmark Model with Political Economy: Link to the Terms-of-Trade Theory

We have thus far assumed that each country's government is benevolent and seeks to maximize the aggregate welfare of its residents. Both casual and formal evidence suggest, however, that it is more realistic to formulate a social welfare function that weights asymmetrically the welfare of different groups in society. The political economy literature has stressed the role of special interest groups in generating these biases in policy (Baron, 1994, Grossman and Helpman, 1996). In this section, we extend the Benchmark Model to allow for government welfare functions that place a higher weight on producer welfare than on consumer welfare. We then consider whether the purpose of a trade agreement in this setting can be given a standard terms-of-trade interpretation along the lines of the (politically augmented) terms-of-trade theories that feature prominently in the trade-agreements literature. For simplicity, except where it might cause confusion we continue to refer to the politically augmented Benchmark Model as simply the Benchmark Model.

4.1 Introducing Political Economy

To represent political-economy motives, we implicitly assume that producers are in a better position to solve the "collective action" problem and hence can better coordinate their demands on the government. We also assume that the ownership of productive assets is highly concentrated, so that we can ignore the role of producers as consumers and as receivers of lump-sum tax rebates. In particular, we let:

$$W^j = CS^j + \gamma^j \pi^j + \text{Trade Tax Revenue}^j, \qquad \text{with } \gamma^j \geq 1, \text{ for } j \in \{H, F\}, \tag{17}$$

where γ^j represents the weight that the government of country j places on the welfare of its producers, with political-economy motives present in country j if and only if $\gamma^j > 1$.

Using (5), (6) and (17), we can write the welfare of the home and foreign governments in the (politically augmented) Benchmark Model as

$$W^H = CS(p_1^H) + \gamma^H \left(\frac{1}{2} \left(1 + \tau_1^H\right) y\left(\hat{x}\right) - \frac{1}{2} \left(\tau_x^H + \tau_x^F\right) \hat{x} \right) + \tau_1^H [D_1(p_1^H) - y\left(\hat{x}\right)] + \tau_x^H \hat{x}$$

and

$$W^F = CS(1) + \gamma^F \left(\frac{1}{2} \left(1 + \tau_1^H\right) y\left(\hat{x}\right) - \frac{1}{2} \left(\tau_x^H + \tau_x^F\right) \hat{x} - \hat{x} \right) + \tau_x^F \hat{x},$$

respectively.[24] Straightforward manipulation of the first-order order conditions that define the Nash policies $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$ delivers:

$$\hat{\tau}_1^{HN} = \frac{\frac{1}{2}\left(\gamma^H - 1\right) y\left(\hat{x}\right) + \frac{1}{2}\hat{x}[y'\left(\hat{x}\right) - \frac{y(\hat{x})}{\hat{x}}]}{-\partial D_1/\partial p_1^H}. \tag{18}$$

Notice that for low enough $\gamma^H > 1$, the home government continues to find it optimal in the Nash equilibrium to set a positive export tax (or import subsidy) on the final good. Nevertheless, when the weight that the home government places on producer surplus becomes sufficiently high, $\hat{\tau}_1^{HN}$ flips sign according to (10) and (18) and becomes positive. In such a case, the home government puts in place a Nash trade policy that leads to an *increase* in the domestic price of the final good (i.e., an import tariff or export subsidy). As we have shown above, these policies tend to transfer surplus from the home country to the foreign country, but a sufficiently politically influenced home government is willing to allow this because consumers bear a disproportionate part of the cost of this rent-dissipation.

Manipulation of the first-order conditions also delivers

$$\hat{\tau}_x^{FN} = \frac{\left(\frac{1}{2}\gamma^F - 1\right)\hat{x}}{\partial \hat{x}/\partial \tau_x^F} < 0,$$

which indicates that for large enough γ^F (in particular $\gamma^F > 2$), the foreign government no longer sets an export tax in the Nash equilibrium but rather chooses to subsidize exports of intermediate inputs. Intuitively, although a subsidy reduces foreign tariff revenue by an amount which is strictly larger than the amount by which foreign profits increase, a sufficiently politically influenced foreign government weights the latter effect disproportionately more, and thus sets a positive export subsidy in the Nash equilibrium.

The fact that the magnitude and even the sign of Nash policies are sensitive to political economy considerations is not particularly surprising: analogous findings are reported for example in Grossman and Helpman (1994) and Bagwell and Staiger (2002, Chapter 10). Nevertheless, as we next demonstrate, the introduction of political economy concerns into the Benchmark Model accentuates its differences with the terms-of-trade theory.

4.2 Alternative Representation of Government Preferences

To facilitate comparison of the Benchmark Model with the terms-of-trade theory, it is helpful to re-express the government objectives contained in (17) as functions of local and "world/international" prices. To this end, we begin by defining the *international price* of the input x, which we denote by p_x^*. In words, p_x^* is the (untaxed) price negotiated in *stage* 3 for the exchange of inputs between the foreign supplier and the home producer. It is easy to see that in the Benchmark Model this

[24] It is straightforward to show that our introduction of political economy into the Benchmark Model does not create a reason for F to utilize τ_1^F, and so we continue to set $p_1^F = 1$ and focus only on the choices of τ_1^H, τ_x^H and τ_x^F.

price is given by $p_x^* = \pi^F/\hat{x} + (1+\tau_x^F)$, which we can write as

$$p_x^* = p_x^*(\hat{x}, \tau_1^H, \tau_x^H, \tau_x^F) \equiv \frac{1}{2}(1+\tau_1^H)\frac{y(\hat{x})}{\hat{x}} - \frac{1}{2}\left(\tau_x^H - \tau_x^F\right). \tag{19}$$

The price of x in the home and foreign country are then given by

$$p_x^H = p_x^H(\tau_x^H, p_x^*) \equiv p_x^*(\hat{x}, \tau_1^H, \tau_x^H) + \tau_x^H \tag{20}$$

and

$$p_x^F = p_x^F(\tau_x^F, p_x^*) \equiv p_x^*(\hat{x}, \tau_1^H, \tau_x^H) - \tau_x^F, \tag{21}$$

respectively. Notice also that $p_x^H - p_x^F = \tau_x^H + \tau_x^F$, and so together with $p_1^H = p_1^H(\tau_1^H) \equiv 1 + \tau_1^H$ we may then re-write (14) as

$$\frac{1}{2}p_1^H y'(\hat{x}) = 1 + \frac{1}{2}\left(p_x^H - p_x^F\right), \tag{22}$$

which implicitly defines

$$\hat{x} \equiv \hat{x}\left(p_1^H(\tau_1^H), p_x^H(\tau_x^H, p_x^*), p_x^F(\tau_x^F, p_x^*)\right). \tag{23}$$

With these definitions, we are now ready to express the government objectives contained in (17) as functions of local and international prices. In particular, using (19) through (23), we have

$$\begin{aligned} W^H &= CS(p_1^H) + \gamma^H[p_1^H y(\hat{x}) - p_x^H \hat{x}] + (p_1^H - 1)[D(p_1^H) - y(\hat{x})] + (p_x^H - p_x^*)\hat{x} \\ &\equiv \bar{W}^H\left(p_1^H(\tau_1^H),\ p_x^H(\tau_x^H, p_x^*),\ p_x^*(\hat{x}, \tau_1^H, \tau_x^H),\ p_x^F(\tau_x^F, p_x^*)\right), \end{aligned}$$

and

$$\begin{aligned} W^F &= CS(1) + \gamma^F \hat{x}[p_x^F - 1] + (p_x^* - p_x^F)\hat{x} \\ &\equiv \bar{W}^F\left(p_x^F(\tau_x^F, p_x^*),\ p_x^*(\hat{x}, \tau_1^H, \tau_x^H, \tau_x^F),\ p_1^H(\tau_1^H),\ p_x^H(\tau_x^H, p_x^*)\right). \end{aligned}$$

Here and throughout this section, we use $\bar{W}^j$ to represent the objectives of government j when expressed as a function *only* of prices.

Notice that, with subscripts on government objective functions denoting partial derivatives, equation (23) implies that the partials $\bar{W}^H_{p_x^*}$ and $\bar{W}^F_{p_x^*}$ are given by $\bar{W}^H_{p_x^*} = W^H_{p_x^*} = -\hat{x}$ and $\bar{W}^F_{p_x^*} = W^F_{p_x^*} = \hat{x}$, and so $\bar{W}^H_{p_x^*} + \bar{W}^F_{p_x^*} = 0$.[25] This reflects the fact that the income effect of the pure terms-of-trade change embodied in the movement of p_x^* – holding local prices and $\hat{x}$ fixed – is given simply by the trade volume ($\hat{x}$). We will return to this property at later points in the section.

4.3 Efficient and Nash Policies

Using these welfare functions, it is straightforward to verify that the efficient policy choices of the two governments (i.e., the policies τ_x^F, τ_x^H and τ_1^H that maximize the sum of home and foreign

[25] In particular, notice that changes in the international price p_x^* have no effect on the choice of x whenever local prices p_x^H and p_x^F are held fixed.

welfare when evaluated in light of the objectives of the governments) must satisfy the following two conditions:[26]

$$\bar{W}^H_{p^H_x} + \bar{W}^F_{p^H_x} + \left[\bar{W}^H_{p^H_x} + \bar{W}^F_{p^H_x} + \bar{W}^F_{p^F_x} + \bar{W}^H_{p^F_x}\right]\left(\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}\right) = 0 \tag{24}$$

and

$$\left\{\bar{W}^H_{p^H_1} - \bar{W}^H_{p^H_x} \cdot \frac{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_1} + \frac{\partial p^*_x}{\partial \tau^H_1}}{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}}\right\} + \left\{\bar{W}^F_{p^H_1} - \bar{W}^F_{p^H_x} \cdot \frac{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_1} + \frac{\partial p^*_x}{\partial \tau^H_1}}{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}}\right\} = 0. \tag{25}$$

The interpretation of (24) and (25) is as follows. Let us begin with the second efficiency condition, whose interpretation is somewhat subtle. The first term on the left-hand side of (25) gives the impact on home welfare of changes in the *mix* of τ^H_1 and τ^H_x which hold fixed p^*_x, and hence, with τ^F_x unchanged, hold fixed as well the level of $p^F_x(\tau^F_x, p^*_x)$. Notice, though, that these tariff changes do not hold fixed $p^H_1(\tau^H_1)$ or $p^H_x(\tau^H_x, p^*_x)$ – and hence, as (22) indicates, do not hold fixed $\hat{x}$. This implies that changes in the mix of τ^H_1 and τ^H_x will in general impact foreign welfare even when they hold p^*_x fixed. The second term on the left-hand side gives the impact on foreign welfare of these tariff changes. Hence, the second efficiency condition in (25) says that small changes in the mix of τ^H_1 and τ^H_x (and the implied p^H_1 and p^H_x) that the home country employs to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x must induce home and foreign welfare changes that balance out to zero on the international efficiency frontier. The first efficiency condition (24) then ensures that the sum of τ^H_x and τ^F_x achieves the efficient level of p^F_x, and hence the efficient level of input trade volume in light of the mix of τ^H_1 and τ^H_x that the home country employs to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x.[27]

Consider next the Nash policies. We can manipulate the first-order conditions that determine

[26] Defining the efficiency frontier in this way fits well with the "member-driven" nature of the WTO, and it is the approach to evaluating the performance of trade agreements taken by most of the literature, but it is not the only approach. An alternative (pursued for example by Aghion, Antràs, and Helpman, 2007 and by Ornelas, forthcoming) is to evaluate the performance of trade agreements on the basis of whether or not the agreement guides governments to a point on an efficiency frontier that is defined with regard to a set of preferences that are unrelated to government preferences (e.g., the maximization of real world income). Such an approach would have some potentially interesting implications in the current setting, because politically motivated governments tend to adopt trade policies that promote producer surplus, and this tends to lead to levels of input production that are increasing in the political weight γ. In a world in which input production is too low as a result of hold up, the policy bias of politically motivated governments could therefore have some attractive consequences from the perspective of real world income. In principle, governments with a particular set of political preferences could adopt policies in the Nash equilibrium which deliver efficient levels of input production in the sense of achieving $\hat{x} = x^E$. Nevertheless, in general there is no reason to expect that political pressures would align themselves in this particular way, and so even from this perspective there would as a general matter be inefficiencies in the Nash equilibrium which a trade agreement could correct.

[27] Using the fact that $1 = \partial p^*_x/\partial \tau^F_x - \partial p^*_x/\partial \tau^H_x$, it is straightforward to verify that whenever the first condition (24) holds, the condition ensuring the efficiency of τ^F_x will also be satisfied. Hence, we can think of condition (24) as effectively pinning down the sum $\tau^H_x + \tau^F_x$.

the unilateral optimal choices of τ_1^H, τ_x^H and τ_x^F, and reduce them to the following two conditions:[28]

$$[\bar{W}^H_{p^H_x} + \bar{W}^F_{p^H_x}] + \left[\bar{W}^H_{p^H_x} + \bar{W}^F_{p^H_x} + \bar{W}^F_{p^F_x} + \bar{W}^H_{p^F_x}\right]\left(\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}\right) + \bar{W}^F_{p^*_x} = 0 \tag{26}$$

and

$$\bar{W}^H_{p^H_1} - \bar{W}^H_{p^H_x} \cdot \frac{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_1} + \frac{\partial p^*_x}{\partial \tau^H_1}}{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}} = 0. \tag{27}$$

Comparing expression (26) to the associated efficiency condition in (24), it is apparent that the two expressions are identical except for the addition in (26) of the left-hand-side term $\bar{W}^F_{p^*_x} = \hat{x} > 0$. This implies that the sum $\tau^H_x + \tau^F_x$ is *inefficiently high* (the first-order condition for efficiency is negative at the Nash taxes), and therefore that the Nash level of input trade volume is inefficiently low in light of the mix of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x. This confirms that the first role of a trade agreement as characterized in Proposition 5 – to better help governments solve the international hold-up problem between producers and suppliers and thereby expand input trade to internationally efficient levels – extends to a political economy setting.

Comparing expression (27) to the associated efficiency condition in (25), it is apparent that the two expressions are identical except for the absence in (27) of the left-hand-side term

$$\bar{W}^F_{p^H_1} - \bar{W}^F_{p^H_x} \cdot \frac{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_1} + \frac{\partial p^*_x}{\partial \tau^H_1}}{\frac{\partial p^*_x}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau^H_x} + \frac{\partial p^*_x}{\partial \tau^H_x}}. \tag{28}$$

Evidently, the *mix* of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver its chosen level of p^*_x and (with τ^F_x fixed) p^F_x is *also inefficient*, as it fails to take account of the fact that the associated movements in p^H_1 and p^H_x – and the changes in $\hat{x}$ that are implied – alter foreign welfare.[29] This confirms that the second role of a trade agreement as characterized in Proposition 5 – to prevent the home government from distorting its final-good market in an effort to extract bargaining surplus from foreign firms – extends to a political economy setting as well.

Consider now the Nash inefficiencies identified by the terms-of-trade theory of trade agreements. According to the broad predictions of the (politically augmented) terms-of-trade theory, an inefficiency akin to the first Nash inefficiency identified above is present, but the second is not (see Bagwell and Staiger, 2001). In the Appendix, we confirm this directly by developing a perfectly competitive version of our (politically augmented) model of input trade, and showing that the predictions of the terms-of-trade theory hold in that setting. Hence we may state:

Proposition 6 *Relative to the terms-of-trade theory of trade agreements, the Benchmark Model*

[28] In order to derive (26), we use the fact that $\bar{W}^H_{p^*_x} + \bar{W}^F_{p^*_x} = 0$, $\partial p^*_x/\partial \tau^H_x = \partial p^*_x/\partial \tau^F_x - 1$, and also $\partial \hat{x}/\partial \tau^H_x = \partial \hat{x}/\partial \tau^F_x$. These properties continue to hold in the model with more general bargaining features developed in section 5.

[29] It is straightforward to show that the additional term in (28) is non-zero under Nash policies whether or not political economy motives are present.

predicts a novel international inefficiency associated with the Nash equilibrium: the mix of policies employed by the home government to deliver its preferred international input price is inefficient.

Intuitively, the difference across the two theories identified in Proposition 6 can be traced to the distinct manner in which the international input price is determined in each theory. According to the terms-of-trade theory, international prices are determined by international market-clearing conditions, and so when the home government chooses the mix of home policies with which to deliver its preferred international input price in the Nash equilibrium, the equilibrium traded input volume ($\hat{x}$) is also held fixed, and therefore the foreign government is indifferent to the home government's chosen policy mix: for this reason, the p_x^*-preserving policy mix chosen by the home government in the Nash equilibrium is clearly efficient from an international perspective. By contrast, according to the Benchmark Model, international (input) prices are determined by bilateral bargaining, and so when the home government chooses the mix of home policies with which to deliver its preferred international input price in the Nash equilibrium, the equilibrium traded input volume ($\hat{x}$) is *not* held fixed, as (22) indicates, and therefore the foreign government is *not* indifferent to the home government's chosen policy mix: for this reason, the p_x^*-preserving policy mix chosen by the home government in the Nash equilibrium is clearly inefficient from an international perspective.

4.4 Politically Optimal Policies

To further interpret the nature of the international inefficiencies exhibited by the Nash policy choices in the Benchmark Model, and in particular to draw comparisons with the terms-of-trade theory concerning the underlying *source(s)* of these inefficiencies, we next follow Bagwell and Staiger (1999) and consider the *politically optimal* tariffs that would arise in the hypothetical situation that governments are *not motivated* by the impact of their tariffs on p_x^* when they make their unilateral tariff choices. In particular, we define

Definition 1 *The home and foreign governments are not motivated by the impact of their tariffs on p_x^* whenever*

$$\begin{aligned}
\bar{W}^H_{p_x^*} \cdot \left(\frac{\partial p_x^*}{\partial \hat{x}} \frac{\partial \hat{x}}{\partial \tau_1^H} + \frac{\partial p_x^*}{\partial \tau_1^H} \right) &= \bar{W}^H_{p_x^*} \cdot \left(\frac{\partial p_x^*}{\partial \hat{x}} \frac{\partial \hat{x}}{\partial \tau_x^H} + \frac{\partial p_x^*}{\partial \tau_x^H} \right) \equiv 0, \ \textit{and} \\
\bar{W}^F_{p_x^*} \cdot \left(\frac{\partial p_x^*}{\partial \hat{x}} \frac{\partial \hat{x}}{\partial \tau_x^F} + \frac{\partial p_x^*}{\partial \tau_x^F} \right) &\equiv 0.
\end{aligned}$$

We next identify the tariffs that would be chosen unilaterally (i.e., non-cooperatively) by governments under this hypothetical behavior and ask whether these tariffs are efficient with respect to actual government preferences. This thought experiment provides a way to define the unilateral choices that governments would make if their countries were "small" in international markets, and by evaluating the efficiency properties of these choices we may thereby assess whether or not the source of Nash inefficiency can be traced to a terms-of-trade driven Prisoners' Dilemma.

Denoting the resulting politically optimal tariffs as $\tau_1^{H,PO}$, $\tau_x^{H,PO}$ and $\tau_x^{F,PO}$, our first result is that the sum $\tau_x^{H,PO}+\tau_x^{F,PO}$ is indeed efficient. Intuitively, this follows from the fact that Nash tariffs on inputs are inefficient because of the extra term $\bar{W}^F_{p_x^*}$ that appears in equation (26); but according to our Definition 1, this term is equal to zero whenever government policies are politically optimal. This indicates that, in the Benchmark Model, it is the terms-of-trade externality that creates a problem (for the level of input tariffs) when governments set tariffs unilaterally.

The result that politically optimal tariffs on input trade are efficient resonates with the results obtained in terms-of-trade theory, but there is an important and subtle difference. In particular, notice that we have defined the hypothetical government behavior associated with the political optimum in a particular way according to Definition 1. An alternative would have been to assume that the home and foreign governments completely *ignore* the impacts of their tariffs on p_x^* when they make their unilateral tariff choices, and therefore act as if

$$\frac{\partial p_x^*}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau_1^H}+\frac{\partial p_x^*}{\partial \tau_1^H}=\frac{\partial p_x^*}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau_x^H}+\frac{\partial p_x^*}{\partial \tau_x^H}\equiv 0 \text{ and } \frac{\partial p_x^*}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau_x^F}+\frac{\partial p_x^*}{\partial \tau_x^F}\equiv 0.$$

It is easy to confirm using (24) and (26) that this alternative behavioral assumption would *not* induce governments to make unilateral input-tariff choices which are consistent with international efficiency.[30]

Putting these observations together, it is now apparent that the Nash inefficiency in input tariffs can be interpreted as a terms-of-trade problem if and only if the "bad" terms-of-trade manipulation – that is, the terms-of-trade manipulation that reflects the pursuit of pure international rent-shifting and is associated with the terms in Definition 1 – is separated from the "good" terms-of-trade manipulation – that is, the terms-of-trade manipulation inherent in the internationally efficient subsidies to input trade and associated with the terms involving $\left(\frac{\partial p_x^*}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau_x^H}+\frac{\partial p_x^*}{\partial \tau_x^H}\right)$ in the Nash optimality condition (26). Nash policies are thus optimal only when the "bad" terms-of-trade manipulation is eliminated, while the "good" terms-of-trade manipulation is preserved. This distinction, however, is immaterial according to the broad predictions of the (politically augmented) terms-of-trade theory;[31] and in the Appendix, we confirm this directly for the perfectly competitive version of our (politically augmented) model of input trade. Hence we may state:

Proposition 7 *In the Benchmark Model, the international inefficiency associated with the Nash choices of input tariffs can be given a modified terms-of-trade interpretation: with regard to input tariffs, the role of a trade agreement is to eliminate the "bad" terms-of-trade manipulation associated with the pursuit of pure international rent-shifting while maintaining the "good" terms-of-trade manipulation inherent in the internationally efficient subsidies to input trade.*

[30] In fact, there is one exception to this statement: in the absence of political economy motives, either of these behavioral assumptions works to induce governments to make unilateral input-tariff choices which are consistent with international efficiency. See also note 34.

[31] While this distinction is immaterial according to the terms-of-trade theory, the political optimum is defined by Bagwell and Staiger (1999) according to the analogue of our Definition 1.

We next assess whether the politically optimal choice of τ_1^H is also efficient. Remember that the Nash equilibrium policy was inefficient because of the extra term in (28). This means that the politically optimal, p_x^*-preserving mix of $\tau_1^{H,PO}$ and $\tau_x^{H,PO}$ satisfies the condition for international efficiency if and only if

$$\bar{W}^F_{p_1^H} - \bar{W}^F_{p_x^H} \cdot \frac{\frac{\partial p_x^*}{\partial \hat{x}} \frac{\partial \hat{x}}{\partial \tau_1^H} + \frac{\partial p_x^*}{\partial \tau_1^H}}{\frac{\partial p_x^*}{\partial \hat{x}} \frac{\partial \hat{x}}{\partial \tau_x^H} + \frac{\partial p_x^*}{\partial \tau_x^H}} = 0. \tag{29}$$

Our definition of political optimality does not generally ensure that this condition will be met. In fact, equation (29) will hold, and thus politically optimal tariffs will be efficient, *only* in the absence of foreign political motives ($\gamma^F = 1$). It is instructive to illustrate this last point more formally. Notice on the one hand, that when $\gamma^F = 1$, we have $W^F = CS(1) + (p_x^* - 1)\,\hat{x}$, and thus $\bar{W}^F_{p_1^H} = (p_x^* - 1)\frac{\partial \hat{x}}{\partial p_1^H}$, $\bar{W}^F_{p_x^H} = (p_x^* - 1)\frac{\partial \hat{x}}{\partial p_x^H}$, and also $\bar{W}^F_{p_x^F} = (p_x^* - 1)\frac{\partial \hat{x}}{\partial p_x^F}$. Because equation (22) implies that $\partial \hat{x}/\partial p_x^H = -\partial \hat{x}/\partial p_x^F$, we can conclude that $\bar{W}^F_{p_x^H} + \bar{W}^F_{p_x^H} = 0$ whenever foreign political motives are absent. On the other hand, the Nash equilibrium choice of τ_x^F imposes the optimality condition

$$-\bar{W}^F_{p_x^F} + \left[\bar{W}^F_{p_x^F} + \bar{W}^F_{p_x^*} + \bar{W}^F_{p_x^H}\right]\left(\frac{\partial p_x^*}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \tau_x^F} + \frac{\partial p_x^*}{\partial \tau_x^F}\right) = 0.$$

Now applying the politically optimal motives described in Definition 1 to this Nash first-order condition together with $\bar{W}^F_{p_x^H} + \bar{W}^F_{p_x^H} = 0$, we immediately obtain $\bar{W}^F_{p_x^F} = 0$, which in turn implies $p_x^* = 1$ and thus $\bar{W}^F_{p_1^H} = \bar{W}^F_{p_x^H} = 0$. Hence, whenever $\gamma^F = 1$, we have that equation (28) will be satisfied and thus the politically optimal mix of $\tau_1^{H,PO}$ and $\tau_x^{H,PO}$ satisfies the condition for international efficiency.

Nevertheless, when foreign political motives are present ($\gamma^F > 1$), the above argument fails and the politically optimal mix of $\tau_1^{H,PO}$ and $\tau_x^{H,PO}$ violates the condition for international efficiency. Hence we may state:

Proposition 8 *In the Benchmark Model, the international inefficiency in the Nash equilibrium associated with the mix of policies employed by the home government to deliver its preferred international input price can be given a modified terms-of-trade interpretation if and only if the foreign government is not motivated by political economy considerations.*

To develop some intuition for the result of Propositions 7 and 8, it is helpful to begin by noting that the efficient choice of each policy must take account of the effects that small adjustments in that policy have on the sum of home and foreign welfare. By definition, when a government sets a policy at its politically optimal level, it ignores the effects of small adjustments in that policy on the welfare of its trading partner, but it also ignores any impacts on its *own* welfare that are generated directly from changes in p_x^*.

From this backdrop, Why are $\tau_x^{H,PO}$ and $\tau_x^{F,PO}$ efficient? Consider $\tau_x^{H,PO}$. When the home government makes a small adjustment to τ_x^H, there are three impacts on foreign welfare $\bar{W}^F$: the effect on foreign welfare of the induced change in p_x^* that runs directly through p_x^*; the effect on foreign welfare of the change in p_x^F associated with the induced change in p_x^*; and the effect on foreign welfare of the (total) induced change in p_x^H. The home government ignores the first effect, but according to the political optimum it ignores this effect on its own welfare as well, and as the two effects cancel out in terms of overall world welfare (recall that $\bar{W}^H_{p_x^*} + \bar{W}^F_{p_x^*} = 0$), no inefficiency arises from this source. The home government also ignores the second and third effects of its adjustment in τ_x^H on foreign welfare, but these effects are identical to the effects on foreign welfare that the foreign government considers when it sets its politically optimal level of τ_x^F, and so the first-order condition that defines $\tau_x^{F,PO}$ ensures that no inefficiency arises from the second and third effects either. A similar logic explains why $\tau_x^{F,PO}$ is also efficient.[32]

Finally, Why does $\tau_1^{H,PO}$ generally fail to be efficient? The reason is that, in contrast to the case of τ_x^H, the foreign government does not possess a policy that can replicate the effects on foreign welfare of τ_1^H at the political optimum, namely, the combined effects of (i) the change in p_x^F and p_x^H associated with the induced change in p_x^*, and (ii) the change in p_1^H that comes from a small adjustment in τ_1^H; and hence, in general, the first-order condition for the foreign country at the political optimum does not prevent inefficiencies from being associated with these effects of a small adjustment in τ_1^H. The exception occurs when the foreign government has no political economy motives. In this case, the foreign government does not value the redistribution of its economy's surplus toward its input suppliers that an increase in p_x^F implies, and therefore a small adjustment in τ_1^H affects foreign welfare at the political optimum in the same way as a small adjustment in τ_x^F, which is to say not at all when evaluated at τ_x^{FPO}. For this reason, in the absence of foreign political economy concerns, $\tau_1^{H,PO}$ is efficient as well.[33]

Hence, in our model the purpose of a trade agreement becomes more complicated when governments possess political economy motives.[34] In this sense, it could be said that we have formally identified a separate "political externality" for a trade agreement to address, somewhat along the lines described in Ethier (2004).[35] This is in stark contrast to the terms-of-trade theory, where the presence or absence of political economy has no impact on the purpose of a trade agreement. The essential difference between the terms-of-trade theory and the theory we develop here that generates this distinction is related to the result reported in Proposition 6, and can again be traced to the difference across the two theories in the way that international prices are determined. According to the terms-of-trade theory, international prices are determined through market clearing conditions,

[32] A natural question is whether the results of Propositions 7 and 8 would be overturned if we relaxed the assumption of symmetric Nash bargaining. But in fact it is straightforwad to show that all of the results of this section hold under generalized Nash bargaining.

[33] As is reflected in our discussion, it is the *foreign* political economy forces that prevent the politically optimal choice of τ_1^H from being efficient. More generally, however, in the presence of symmetric home-supplier/foreign-producer relationships, political economy forces in either country will interfere with the efficiency properties of the political optimum.

[34] We emphasize this feature in the context of Proposition 8, although one could draw a similar inference from the observation (see note 30) that the distinction between "good" and "bad" terms-of-trade manipulation only becomes relevant in the Hold-up Model once political-economy motives are introduced.

[35] As Ethier (2004, p. 305) puts it, " 'Political externalities,' by my definition, arise when policymakers in one country believe that their political status (whatever that might be specified to mean) is directly sensitive, to some degree, to actions by policymakers in another country." See also Bagwell and Staiger (2002, Chapter 2) for a further discussion of these ideas.

with local prices in each economy then determined by arbitrage conditions that link international prices with local prices through a country's own policies. When countries choose politically optimal policies according to the terms-of-trade theory, they therefore ensure that the local-price effects of small adjustments in the policies of their trading partners can have no first-order effect on their welfare, because these local price effects could already have been generated by their own policy adjustments, and their politically optimal first-order-condition ensures that the welfare impacts are zero. By contrast, in the theory we develop here, international prices are determined by bilateral bargaining between foreign suppliers and home producers, and given the more complex channels through which a trading partner's policies can induce local effects in a country's economy, the country's politically optimal first-order conditions cannot – except for the special case in which governments maximize real national income – ensure that the welfare effects will be zero, because there is no longer any guarantee that the local effects of a trading partner's policy adjustments could have been generated by a country's own policy adjustments.

Of course, both our Benchmark and Benchmark Models impose many strong assumptions, and it is important to know whether the insights developed thus far reflect special features of these strong assumptions or rather are more general. In the next section, we explore this issue.

5 Sensitivity

In this section we consider the generality of our central findings to various alternative modeling assumptions. For simplicity, we return to a setting in which governments do not possess political economy motives.

5.1 General Bargaining Power

In the Benchmark Model we have assumed that all bargaining between home producers and foreign suppliers is characterized by symmetric Nash bargaining, with each party capturing one-half of the surplus. We have also assumed that the outside option of each party is determined in a secondary market located in the home country in which the two parties also have symmetric bargaining power. We now explore alternative assumptions. Below we consider the case of general bargaining power. In the Appendix we also consider the possibility that the secondary market for the foreign supplier is located in the foreign country. In each case, we show that helping governments solve the international hold-up problem between producers and suppliers and avoid the policy distortions that would be introduced by attempts to extract bargaining surplus from foreign firms remain the fundamental purposes of a trade agreement in the model.

Consider first a situation in which the bargaining in both stages 3 and 4 is captured by the generalized Nash bargaining solution with weights α and $(1-\alpha)$ for the home producer and foreign supplier, respectively, where $\alpha \in (0,1)$. As before, we assume that all agents have an ex-ante zero outside option. The sequence of events is otherwise identical to that of the Benchmark Model.

We focus directly on deriving Nash policy choices, assuming as before that the home and foreign governments select their respective tariffs simultaneously in a prior *stage* 0. Following analogous steps as in previous sections, it is easy to see that the final-good producer in H now has a *stage*-3 payoff of $\alpha\left(1+\tau_1^H\right)y\left(x\right)-\alpha\left(\tau_x^H+\tau_x^F\right)x$, with the supplier in F now receiving a *stage*-3 payoff of $\left(1-\alpha\right)\left(1+\tau_1^H\right)y\left(x\right)-\left(1-\alpha\right)\left(\tau_x^H+\tau_x^F\right)x$. As a result, we have that the *stage*-2 choice of $\hat{x}$ is now defined by

$$\left(1-\alpha\right)\left(1+\tau_1^H\right)y'\left(\hat{x}\right)=1+\left(1-\alpha\right)\left(\tau_x^H+\tau_x^F\right), \tag{30}$$

and hence the *stage*-1 payoffs of the home and foreign firm are given by

$$\begin{aligned}\pi^H &= \alpha\left(1+\tau_1^H\right)y\left(\hat{x}\right)-\alpha\left(\tau_x^H+\tau_x^F\right)\hat{x}, \text{ and}\\ \pi^F &= \left(1-\alpha\right)\left(1+\tau_1^H\right)y\left(\hat{x}\right)-\left(1-\alpha\right)\left(\tau_x^H+\tau_x^F\right)\hat{x}-\hat{x}.\end{aligned}$$

Home and foreign welfare may now be written as

$$\begin{aligned}W^H &= CS(p_1^H)+\pi^H+\tau_1^H\left[D_1(p_1^H)-y\left(\hat{x}\right)\right]+\tau_x^H\hat{x}, \text{ and}\\ W^F &= CS(1)+\pi^F+\tau_x^F\hat{x}.\end{aligned} \tag{31}$$

The first-order conditions that define the Nash policies $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$ can be manipulated using (30) to yield[36]

$$\begin{aligned}\hat{\tau}_1^{HN} &= \frac{\left(1-\alpha\right)\hat{x}\left[y'\left(\hat{x}\right)-\frac{y(\hat{x})}{\hat{x}}\right]}{-\partial D_1/\partial p_1^H},\\ \hat{\tau}_x^{HN} &= -\frac{\left[\alpha-\left(1-\alpha\right)\hat{\tau}_1^{HN}\right]y'\left(\hat{x}\right)}{\left(1-\alpha\right)}+\frac{\alpha\hat{\tau}_x^{FN}}{\left(1-\alpha\right)}-\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H}, \text{ and}\\ \hat{\tau}_x^{FN} &= -\alpha\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F}.\end{aligned}$$

By setting $\alpha=1/2$, it is easily verified that these expressions coincide with those of the Benchmark Model above, where symmetric Nash bargaining was assumed. These equilibrium conditions illustrate that as long as home producers have less than full bargaining power (i.e., $\alpha<1$), their will still be an incentive to use trade taxes on the final good to reduce its price in the home market, that is $\hat{\tau}_1^{HN}<0$. The intuition for this is analogous to that in the Benchark Model. Similarly, as long as foreign suppliers have less than full bargaining power (i.e., $\alpha>0$), they will also have an incentive to resort to an export tax, simply because they bear less than the full cost of any increase in their marginal cost of production. Thus the foreign government can pass part of the cost of an export tax on to the home country, and it will choose to do so in a non-cooperative tariff equilibrium.[37]

[36] Throughout the extension section, we simply assume that second-order conditions are met.

[37] It is interesting to note that our Benchmark Model suggests a novel explanation for the observation that developing countries often appear to use trade policy as a means to raise revenue. The standard explanation is that these countries, while small in world markets, impose taxes on trade because their governments have revenue needs but do not have access to more appropriate measures of the kind typically available to developed-country governments for taxing citizens to raise revenue (e.g., an income tax); and so these governments are forced to use trade taxes as a second-best measure for raising revenue from their own citizens. In our Benchmark Model, governments have

It bears re-emphasis here that although the goal of an export tax in our model is similar to that in the traditional terms-of-trade theory, the determinants of the size of this tax are quite distinct. In particular, the traditional literature emphasizes the importance of size (Johnson, 1953, Kennan and Riezman, 1988, Bond and Syropoulos, 1996): larger countries will tend to set higher export taxes. The reason is that in the standard terms-of-trade theory, exporter prices are determined through market clearing conditions, and so a country can alter these prices only if it is big enough to make a difference to the market clearing price when it alters its excess supplies to the market. In our theory, by contrast, exporter prices are determined as a result of bilateral bargaining, and so a country can alter exporter prices if it can alter the conditions of bilateral bargaining that determine those prices. Such a country need not be "big" in the traditional sense, and in fact in our theory a prerequisite for export taxes to be used is for the foreign country suppliers to have weak (or at least less-than-complete) bargaining power in the market for inputs.[38]

Leaving aside the differences that the specifics of the bargaining process make to the equilibrium choices of $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$, it is important to emphasize two general results that continue to hold regardless of the value of α. First, the international hold-up problem persists and the volume of international input trade is inefficiently low as a consequence. To see this, one can manipulate the first-order conditions and use the expression for $\partial\hat{x}/\partial\tau_x^H$ and $\partial\hat{x}/\partial\tau_x^F$ implied by (30) to derive

$$y'(\hat{x}) = 1 - \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H} > 1,$$

which implies that $\hat{x} < x^E$. Second, our model continues to predict that there are distortions in the final good market that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers. The purpose of a trade agreement therefore remains to help governments better solve these two problems.

It is also possible to use the extended model developed in this section to make a broader point.

no revenue needs of their own: rather, they are assumed only to use trade taxes as a way to increase real national income. The foreign country in our Benchmark Model could be interpreted as a less-developed country because its firms produce the inputs that are finished for sale in the home (developed) country. Interpreted in this way, our Benchmark Model then indicates that the use of tariffs by developing-country governments for the purpose of raising revenue may in fact not be a second-best measure for raising revenue from their own citizens to fund public expenditures, but rather represents a first-best measure for extracting revenue from developed-country producers in light of the poor bargaining power of the less-developed-country supplier firms in their bilateral relationship with developed-country producers.

[38] In fact what is key is that foreign suppliers have less than full bargaining power *in the secondary market* for inputs. To see this, consider the case in which there is generalized Nash bargaining in both stages 3 and 4, but with potentially different bargaining weights α_P and α_S, respectively. In such a case, the Nash tariff choices are characterized by the following conditions:

$$\begin{aligned}
\hat{\tau}_1^{HN} &= \frac{(1-\bar{\alpha})\,\hat{x}[y'(\hat{x}) - \frac{y(\hat{x})}{\hat{x}}]}{-\partial D_1/\partial p_1^H}, \\
\hat{\tau}_x^{HN} &= -\frac{\left[\bar{\alpha} - (1-\bar{\alpha})\,\hat{\tau}_1^{HN}\right] y'(\hat{x})}{(1-\alpha_S)} + \frac{\alpha_S \hat{\tau}_x^{FN}}{(1-\alpha_S)} - \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H}, \text{ and} \\
\hat{\tau}_x^{FN} &= -\alpha_S \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F},
\end{aligned}$$

where $\bar{\alpha} \equiv \alpha_S\delta + \alpha_P(1-\delta)$. It is apparent that $\hat{\tau}_x^{FN} > 0$ only if $\alpha_S > 0$.

Up to now we have not taken a stance as to whether the home producer and foreign supplier are vertically related or not. According to the transaction-cost approach to the boundaries of the firm (c.f., Coase, 1936, Williamson, 1985), vertical integration would arise precisely when the hold-up inefficiencies that we have modelled above become large relative to the larger "governance" costs of running an integrated organization. According to that view, our novel rationale for trade agreements would disappear because production of the final good could then be characterized by neoclassical production theory.[39] Nevertheless, the property-rights approach to the theory of the firm (c.f., Grossman and Hart, 1986, Hart and Moore, 1990) has persuasively argued that firm boundaries are better understood as determining the relative bargaining power of producers (via the allocation of residual rights of control inherent in the ownership of productive physical assets) rather than as affecting the space of contracts available to economic agents. Under this interpretation, the rationale for trade agreements that we propose in this paper would very much apply to vertical integrated cross-border production relationships. A crude way to capture the essence of the property-rights theory of the firm in terms of the extended model developed in this section would be to associate international outsourcing relationships with a low value of α (the bargaining power of home producers) as compared to the value of α applying to international insourcing relationships. With this interpretation, our finding that the fundamental purpose of a trade agreement does not depend on the value of α then suggests as well that the presence or absence of vertical integration would not alter the fundamental purpose of a trade agreement.[40]

5.2 Ex-Ante Lump-Sum Transfers

Our Benchmark Model rules out ex-ante lump-sum transfers between home producers and foreign suppliers. Although this seems a plausible assumption in our international framework where the promises associated with these transfers may be hard to enforce, it is useful to study the robustness of our results. In particular, it may be thought that when one of the parties is able to fully internalize the gains from offshoring, the Nash equilibrium policy choices will cease to be inefficient. We next show that this is never the case. For that purpose, we consider the following modification of the stage 0 of our Benchmark Model:

stage 1. The unit measure of producers in H and suppliers in F are randomly matched, producing a unit measure of matches. Each producer in H and its matched supplier in F bargain over whether to continue their relationship or not and lump-sum transfers are allowed in the bargaining. This *stage*-1 bargaining captured by the generalized Nash bargaining solution with weights β and $(1-\beta)$ for the home producer and foreign supplier, respectively, where

[39] This assumes that home producers could hire foreign suppliers in a competitive market at a given price, which is consistent with the transaction-cost assumption of a frictionless integrated structure (see Grossman and Helpman, 2002, for a general equilibrium treatment).

[40] This is not to say that the presence or absence of vertically integrated home producers and foreign suppliers would be irrelevant for the nature of trade agreements. On the contrary, to the extent that international factor ownership associated with vertically integrated multinational firms alters the objective functions of each government, the nature of trade agreements could be very much affected (see Blanchard, 2006). Rather, our point is simply that vertical integration does not by itself obviate the need for a trade agreement to address the international hold-up problem.

$\beta \in (0,1)$. If the relationship is terminated, both firms exit; if an agreement is reached, the producer retains the supplier and provides it with a list of customized input specifications.

For simplicity, we assume that the remaining stages of the game are as in the Benchmark Model (and, in particular, all *ex-post* bargains are governed by symmetric Nash bargaining). This implies that at stage 1, the home producer and the foreign supplier anticipate that if they reach an agreement, they stand to obtain a joint payoff of

$$\pi^H + \pi^F = \left(1 + \tau_1^H\right) y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right) \hat{x} - \hat{x},$$

where $\hat{x}$ is given by

$$\frac{1}{2}(1 + \tau_1^H) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\left(\tau_x^H + \tau_x^F\right).$$

Instead, if an agreement is not reached, both firms exit and are left with a payoff equal to 0. In the Appendix, we show that $\pi^H + \pi^F > 0$, which implies that all pairs reach an agreement at stage 0. Note, however, that because of the lump-sum transfers, the division of profits between home producers and foreign suppliers is now detached from the ex-post bargaining solution.[41] In particular, we have:

$$\begin{aligned}
\pi^H &= \beta\left[\left(1 + \tau_1^H\right) y\left(\hat{x}\right) - (\tau_x^H + \tau_x^F)\hat{x} - \hat{x}\right], \text{ and} \\
\pi^F &= (1-\beta)\left[\left(1 + \tau_1^H\right) y\left(\hat{x}\right) - (\tau_x^H + \tau_x^F)\hat{x} - \hat{x}\right].
\end{aligned}$$

Whenever $\beta = 1$, we have that Home producers have ex-ante full bargaining power, which corresponds to a situation in which they can make take-it-or-leave-it offers to foreign suppliers thus capturing all the surplus. Conversely, the case with $\beta = 0$ is one in which foreign suppliers have all the bargaining power. The values of home and foreign welfare are still given by equation (31) with the new profit levels π^H and π^F applying.

We can next turn to study the Nash equilibrium policy choices of this variant of the model with lump-sum transfers. Manipulating the first-order conditions related to the choices of $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$ delivers:

$$\begin{aligned}
\hat{\tau}_1^{HN} &= \frac{(1-\beta)\,\hat{x}\left[y'\left(\hat{x}\right) - \frac{y(\hat{x})}{x}\right]}{-\partial D_1/\partial p_1^H}, \\
\hat{\tau}_x^{HN} &= -\left(1-\beta\right)\frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H} - \beta + \hat{\tau}_1^{HN} y'\left(\hat{x}\right), \text{ and} \\
\hat{\tau}_x^{FN} &= -\beta\frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^F} - (1-\beta).
\end{aligned}$$

It is instructive to consider first the case in which home producers have full bargaining power, that is $\beta = 1$. In such a case, we obtain that $\hat{\tau}_1^{HN} = 0$ and $\hat{\tau}_x^{HN} = -1$, which indicate that

[41] Still, the equilibrium level of $\hat{x}$ will be identical to that in the Benchmark Model, since foreign suppliers choose $\hat{x}$ to maximize ex-post payoffs (thus ignoring ex-ante payments).

if the foreign government followed a policy of free trade, the Nash equilibrium policies would coincide with the second-best ones and there would not be the need for a trade agreement. The intuition is straightforward: the home government would in such a case be internalizing the whole gains from trade and thus its policy choices would coincide with those of a world social planner facing the same constraints on the choice of $\hat{x}$. The last first-order condition shows however that the foreign government will not remain passive and will in fact set a positive export tax equal to $\hat{\tau}_x^{FN} = -\hat{x}/\left(\partial\hat{x}/\partial\tau_x^F\right) > 0$, which is equivalent to the export tax that maximizes export tax revenue $\tau_x^F\hat{x}$. The logic is as follows: absent government intervention, foreign input suppliers gain nothing from their bilateral relationships with home final good producers when $\beta = 1$, and as a consequence welfare in F is unaffected by the hold-up problem. This gives the foreign government no incentive to intervene to try to fix the problem. Instead, as the foreign government can commit to an export policy before its suppliers are matched with producers in the home country (i.e., before *stage* 1), its only goal is to collect maximal trade tax revenue from the exportation of x, because the incidence of its tax is borne completely by final good producers in H (through a reduction in the transfer that can be extracted from the foreign supplier by the domestic producer).[42]

The opposite case in which foreign suppliers have all the bargaining power ($\beta = 0$) does not produce international efficient policies either. Note that, in such a case, the foreign government sets the efficient subsidy level $\hat{\tau}_x^{FN} = -1$, but the home country ceases to maintain free trade in the final good in an attempt to extract surplus from foreign suppliers. In fact, it is straightforward to show that, regardless of the value of β, the Nash equilibrium policy choices will necessarily lead to an inefficiently low level of input trade. In particular, manipulation of the first-order conditions delivers that, for any $\beta \in (0,1)$,

$$y'(\hat{x}) = 1 - \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F},$$

which again implies that that $\hat{x} < x^E$. Thus, our model generates a rationale for trade agreements even in the presence of ex-ante transfers and independently of the level of ex-ante bargaining power.

5.3 Multiple Foreign Countries and Search Costs

Our analysis has been restricted to situations in which home producers can only search for suppliers in F. It is straightforward to show that at least some of our results could be overturned when this restriction is relaxed. To illustrate this in the starkest possible way, consider again the variant of the Benchmark Model described in the last subsection, with lump-sum transfers available to producers. Assume further that $\beta = 1$, so that home producers have full bargaining power and set $\hat{\tau}_1^{HN} = 0$ and $\hat{\tau}_x^{HN} = -1$. As shown above, international efficiency is not reached because the foreign government has an incentive to set a positive export tax in that case. Now consider the

[42] Naturally, the ability of the foreign government to extract surplus from home producers partially relies on the assumption that home producers can costlessly seek suppliers only in the foreign country. In the next subsection, we develop an extension of the model in which home producers can search for suppliers in one of two foreign countries. We show there that the existence of distortions persist even when one of these countries becomes infinitesimally small in relation to the size of world markets.

case in which there is a second "foreign" country, denoted by S for "South," with an additional unit measure of potential suppliers identical to those in F. Assume that Foreign and South are identical in every other respect, including preferences, technology and bargaining strength. Under these circumstances and as long as $\tau_x^F > \tau_x^S$, all home producers will prefer to match with southern suppliers over suppliers in F. As a result, the government in F will have an incentive to reduce its export tax below the southern one. Pushing this argument further, it is straightforward to show then that the optimal foreign and southern export taxes that emerge from this variant of the model are 0.[43] As a result, in this extended variant of the Benchmark Model, home producers internalize the whole gains from the offshoring relationship, and the home government chooses the first-best policies described in Proposition 2. The rationale for a trade agreement in this extended variant of the Benchmark Model has vanished.

This example, however, is special in a number of ways. To begin with, the assumption that F and S are symmetric is not innocuous: if one of the two foreign countries has a comparative advantage in supplying inputs, it can (and will) maintain a positive export tax (analogous to "limit pricing" in the case of Bertrand competition among firms), and the result of our Benchmark Model is then preserved. More importantly, the structure of the example above imposes that home producers find a match with probability one, no matter where they search for suppliers. As emphasized by Grossman and Helpman (2005), an important feature of offshoring relationships is the costly search for suitable partners. The same characteristics that make offshoring relationships contractually difficult (i.e., customization, international enforceability of contracts, etc.) preclude the existence of a frictionless competitive market for inputs or for suppliers.

To illustrate these issues in a simple way, we now assume that F contains a measure ρ of potential suppliers, while S contains a measure $1-\rho$. If k home producers search for matches in F, the total number of successful matches there is given by the matching function $m(k,\rho) \leq \min\{k,\rho\}$, where $m(\cdot)$ is increasing in both arguments and features constant returns to scale. For simplicity, we assume that S adopt a laissez-faire policy. Will this force F to give up the use of an export tax? As we now demonstrate, the answer is "No."

To show this, we begin by noting that, for home producers to be indifferent between searching in F and in S, we need:

$$\frac{m(k,\rho)}{k}\left(y\left(\hat{x}^F\right) - \left(1+\tau_x^F\right)\hat{x}^F\right) = \frac{m(1-k,1-\rho)}{1-k}\left(y\left(\hat{x}^S\right) - \hat{x}^S\right), \tag{32}$$

where $\hat{x}^F$ is such that $\frac{1}{2}y'\left(\hat{x}^F\right) = 1+\tau_x^F+\tau_x^H$, while $\hat{x}^S$ is such that $\frac{1}{2}y'\left(\hat{x}^S\right) = 1+\tau_x^H$. Equation (32) defines a negative relationship between k and τ_x^F: intuitively, an increase in the foreign export tax should be matched by an increase in the probability of finding a match in that country, which in turn requires a decrease in the measure of home producers searching for partners in that country. To see this formally, note that using the assumption of constant-return-to scale in the matching

[43] The logic is analogous to that behind the fact that Bertrand competition implies marginal-cost pricing.

function, we can write:

$$\frac{dk}{d\tau_x^F} = \frac{\left(y'\left(\hat{x}^F\right) - \left(1+\tau_x^F\right)\right)\left(-\partial\hat{x}^F/\partial\tau_x^F\right) + \hat{x}^F}{-\frac{\rho\mu'(\rho/k)}{k\mu(\rho/k)}\frac{1}{k}\left(y\left(\hat{x}^F\right) - \left(1+\tau_x^F\right)\hat{x}^F\right) - \frac{1}{(1-k)^2}\frac{(1-\rho)\mu'((1-\rho)/(1-k))}{\mu(\rho/k)}\left(y\left(\hat{x}^S\right) - \hat{x}^S\right)} < 0, \quad (33)$$

where $\mu(\rho/k) \equiv m(1, \rho/k)$.

In order to explore the implications of this framework for the optimal choice of an export tax in F, we first define welfare in F as the sum of consumer surplus and tariff revenue collected from all the matched bilateral pairs:

$$W^F = CS(1) + m(k,\rho)\,\tau_x^F \hat{x}^F,$$

It thus follows that optimal choice of τ_x^F (denoted $\hat{\tau}_x^F$) will now satisfy:

$$\frac{\partial W^F}{\partial \tau_x^F} = \frac{\partial m(k,\rho)}{\partial k}\frac{dk}{d\tau_x^F}\tau_x^F\hat{x}^F + m(k,\rho)\,\hat{x}^F + m(k,\rho)\,\tau_x^F\frac{\partial \hat{x}^F}{\partial \tau_x^F} = 0,$$

which in turn implies:

$$\hat{\tau}_x^F = \frac{\hat{x}^F}{-\frac{\partial \hat{x}^F}{\partial \tau_x^F} - \frac{\partial m(k,\rho)}{\partial k}\frac{1}{m(k,\rho)}\frac{dk}{d\tau_x^F}\hat{x}^F} > 0.$$

In sum, provided that $\frac{dk}{d\tau_x^F}$ remains bounded, the optimal export tax will be positive. It is straightforward to show that for well-behaved matching functions, the export tax will remain positive even for infinitesimally small countries. In particular, notice from equation (32) that whenever the elasticity of $m(\cdot)$ with respect to both of its arguments is positive, we will have that when $\rho \to 0$, and hence as F becomes infinitesimally small, $\frac{dk}{d\tau_x^F}$ goes to 0 as well, and thus

$$\hat{\tau}_x^F \to \frac{\hat{x}^F}{-\partial\hat{x}^F/\partial\tau_x^F},$$

which is the expression we derived in the previous section when $\beta = 1$ and only Foreign was the source of inputs.[44]

5.4 Ad Valorem Tariffs

We next return to the Benchmark Model but depart from the specific tariff analyzed in the previous sections, and consider instead an ad-valorem import tariff on intermediate inputs. We show here that ad valorem tariffs introduce a novel channel through which bargaining between the home producer and foreign supplier can be affected. Despite this novel channel, however, we confirm that the role played by an international trade agreement remains the same.

To this end, with the "international" (foreign exporter) price p_x^* still denoting the price negoti-

[44] In particular, under the maintained assumptions, $\frac{\rho\mu'(\rho/k)}{k\mu(\rho/k)}$ is positive and bounded below 1, while $\frac{(1-\rho)\mu'((1-\rho)/(1-k))}{\mu(\rho/k)}$ goes to infinity when $\rho \to 0$.

ated in *stage* 3 for the exchange of intermediate inputs between the foreign supplier and the home producer, we now let t_x^H and t_x^F denote, respectively, the home-country and foreign-country taxes on trade in the intermediate good x expressed in ad valorem terms. With this notation we highlight explicitly that the *stage*-3 negotiation between producer and supplier divides surplus between them by agreeing on the *price* at which the foreign supplier sells the x units of intermediate input to the home producer. To focus on the novel aspects of ad valorem tariffs, we now ignore tariffs on the final good and set $\tau_1^H = \tau_1^F \equiv 0$. With this assumption, according to the Benchmark Model there would be only one problem for a trade agreement to solve in the presence of specific tariffs on trade in the intermediate input x, namely, the elimination of the international hold-up problem, and we now confirm that this remains the case when the tariffs take an ad valorem form.

Specifically, if the producer and supplier reach an agreement in their *stage*-3 bargaining that specifies a price level $\tilde{p}_x^*$, then the home-country producer receives a *stage*-3 payoff of $\omega^H = y(x) - (1+t_x^H)\tilde{p}_x^* x$ while the foreign supplier receives a *stage*-3 payoff of $\omega^F = \frac{\tilde{p}_x^*}{(1+t_x^F)}x$. Notice that this implies a bargaining frontier defined by $\omega^H = y(x) - (1+t_x^H)(1+t_x^F)\omega^F$: because the level of the exporter price p_x^* is used by the home producer and foreign supplier to shift surplus between them, a positive ad valorem import tariff or export tax makes the *slope* of the bargaining frontier between the home producer and the foreign supplier steeper, while a negative ad valorem tariff (an import or export subsidy) makes the slope of the bargaining frontier flatter.[45] In effect, then, ad valorem trade taxes penalize the producer and supplier for shifting surplus toward the foreign supplier (with a high p_x^*), while ad valorem trade subsidies encourage surplus-shifting in this direction, suggesting a novel channel through which ad valorem trade taxes can affect the severity of the international hold-up problem. This channel is not present when a specific tariff is instead utilized, because the slope of the bargaining frontier between producer and supplier is -1 independent of the level of the specific tariffs τ_x^H and τ_x^F.

On the other hand, if the producer and supplier fail to reach an agreement in their *stage*-3 bargaining, then they will be forced to bargaining with different agents in the secondary market in stage 4. At that point the outside options are zero, while sale revenue is $\delta y(x)$. The *stage*-4 Nash bargaining then solves $max_{\omega_S^H,\omega_S^F}\ \omega_S^H \cdot \omega_S^F$ subject to $\omega_S^H = \delta y(x) - (1+t_x^H)(1+t_x^F)\omega_S^F$. The resulting *stage*-3 outside options for the home producer and foreign supplier are $\frac{1}{2}\delta y(x)$ and $\frac{1}{2}\delta y(x)/[(1+t_x^H)(1+t_x^F)]$, respectively.

The *stage*-3 Nash bargaining problem between the home producer and foreign supplier can now

[45] We abstract here from the possibility that firms might engage in transfer-pricing-type behavior in order to avoid trade taxes or collect trade subsidies. In our setting, this amounts to assuming that firms do not have other (non-price) means to transfer surplus between them in their bilateral bargain. If they did have such means, then the price they negotiate would be determined completely by the sign of the trade taxes subject only to the ability of governments to regulate such behavior. Even without such means, the firms in our model do respond to government trade taxes by negotiating different prices, but at least when these firms are taken to be engaged in arms-length transactions this would not be interpreted as transfer pricing in the traditional sense.

be characterized as follows:

$$\begin{aligned}&Max_{\omega^H,\omega^F}\left(\omega^H-\frac{1}{2}\delta y(x)\right)\left(\omega^F-\frac{1}{2}\delta y(x)/[(1+t_x^H)(1+t_x^F)]\right)\\&s.t.\quad \omega^H=y(x)-(1+t_x^H)(1+t_x^F)\omega^F.\end{aligned}$$

The solution to this bargaining problem yields $\omega^H=\frac{1}{2}y(x)$ and $\omega^F=\frac{1}{2}\frac{y(x)}{(1+t_x^H)(1+t_x^F)}$, and an implied foreign exporter price of $\hat{p}_x^*=\frac{\frac{1}{2}y(x)}{(1+t_x^H)x}$. The choice of x at *stage* 2 is then governed by

$$\frac{1}{2}y'(\hat{x})=(1+t_x^H)(1+t_x^F), \tag{34}$$

and hence $\hat{x}$ continues to be decreasing in the (ad valorem) tariffs t_x^H and t_x^F, despite the novel channel through which the ad valorem tariffs affect the bargaining between home producer and foreign supplier. With this, we can now write the *stage*-1 payoffs of the home and foreign firm as

$$\begin{aligned}\pi^H&=\frac{1}{2}y(\hat{x}),\ \text{and}\\ \pi^F&=\frac{1}{2}\frac{y(\hat{x})}{(1+t_x^H)(1+t_x^F)}-\hat{x}.\end{aligned}$$

We consider next the Nash tariff choices. With $\tau_1^H=\tau_1^F\equiv 0$, home and foreign welfare are now given by

$$\begin{aligned}W^H&=CS(1)+\pi^H+\tau_x^H\hat{p}_x^*\hat{x}=\frac{1}{2}y(\hat{x})+\frac{1}{2}\frac{\tau_x^H y(\hat{x})}{(1+t_x^H)},\\ W^F&=CS(1)+\pi^F+\frac{\tau_x^F}{1+\tau_x^F}\hat{p}_x^*\hat{x}=\frac{1}{2}\frac{y(\hat{x})}{(1+t_x^H)}-\hat{x}.\end{aligned}$$

Using the expression for $\hat{x}$ in (34), it is direct to show that the first order condition for t_x^F implies $\hat{t}_x^{FN}=0$: as was the case with specific tariffs, when tariffs take an ad valorem form the foreign country has no unilateral incentive to intervene in the Benchmark Model.

To check whether $\hat{t}_x^{HN}$ might achieve international efficiency in light of $\hat{t}_x^{FN}=0$, we make use of (34) to observe that, in combination with $t_x^F=0$, international efficiency would require $t_x^H=-\frac{1}{2}$. But differentiating W^H with respect to τ_x^H and evaluating at $t_x^F=0$ yields

$$\frac{\partial W^H}{\partial \tau_x^H}=\frac{1}{2}y'(\hat{x})\frac{\partial \hat{x}}{\partial t_x^H}+\frac{1}{2}\frac{y(\hat{x})+\tau_x^H y'(\hat{x})\frac{\partial \hat{x}}{\partial t_x^H}}{(1+t_x^H)}-\frac{1}{2}\frac{\tau_x^H y(\hat{x})}{(1+t_x^H)^2},$$

which is strictly positive when evaluated at the internationally efficient level of $t_x^H=-\frac{1}{2}$: by implication, then, $\hat{t}_x^{HN}$ is higher than the internationally efficient level.

Hence, while the mechanisms through which specific and ad valorem tariffs on traded inputs influence the international hold-up problem are distinct, the broad conclusions are the same. Under Nash policy choices, the international hold-up problem persists, the volume of international input trade is inefficiently low as a consequence, and the purpose of a trade agreement (when $\tau_1^H = \tau_1^F \equiv 0$) remains to help governments better solve the problem.[46]

6 Conclusion

In this paper, we have initiated the study of trade agreements in the presence of offshoring. Our findings suggest that a number of the salient features of trade in inputs – including the prominent role played by relationship-specific investments and the associated contracting difficulties that are likely to give rise to international hold-up problems – have important implications for the nature and purpose of trade agreements. If governments seek to maximize real national income, then the purpose of a trade agreement in this environment is to induce governments to solve the international hold-up problem by exploiting their ability to affect international prices with their trade policy choices, while at the same time preventing governments from seeking to exploit their power over international prices for traditional terms-of-trade reasons. If governments are instead motivated by political economy/distributional concerns, the purpose of a trade agreement becomes more complex, and cannot be reduced to solving a simple terms-of-trade problem. Finally, regardless of the objectives of governments, the degree of inefficiency in the unilateral trade policy choices of governments depends on the details of the ex-post producer-supplier bargaining environment, and while input trade must always be increased in an efficient international agreement, efficiency may require that final goods trade is diminished. As we have observed, these findings are at odds with existing theories of trade agreements, and they suggest a novel interpretation of the value of trade agreements when offshoring is prevalent.

Our paper raises many new questions, both theoretical and empirical. Are international prices best thought of as determined through countless bilateral bargains between buyers and sellers, or rather through anonymous market clearing mechanisms? To the extent that it is the former, Do the trade policy stances of governments in practice have systematic impacts on bargaining outcomes and, through this channel, on trade volumes? To what extent can the architecture of the WTO, including its emphasis on reciprocity and non-discrimination, be understood from the perspective of the theory we develop here? Does the changing nature of international trade indicate the need for fundamental changes in the international institutions that govern the world trading system? These and related questions strike us as especially fertile areas for further research.

[46] It is interesting to observe that the novel channel through which ad valorem tariffs alter the bargaining outcome between home producer and foreign supplier – namely, the slope of the bargaining frontier – also suggests that these policy instruments may have a broader class of applicability with regard to their ability to mitigate international hold-up problems than is the case for specific tariffs. For example, if x were reinterpreted as the unverifiable quality of a fixed unit to be traded, so that tariff policy could not then be conditioned on x, a specific tariff on trade in x would loose its ability to affect the hold-up problem, but an ad valorem tariff would continue to be useful in this regard.

Appendix A

A.1. Non-Negativity Constraints

In the main text, we have ignored situations in which equilibrium trade policies might violate the non-negativity constraints on the outside options and the surplus available to agents in the negotiation. The purpose of this Appendix is to explore those situations and show how they do not invalidate the main results of the paper. To save space, we focus on an analysis of the Benchmark model, and only briefly comment on the model with generalized bargaining developed in section 5.1. In the Benchmark model, the surplus over which the producer and the supplier bargain is given by

$$\left(1+\tau_1^H\right) y\left(\hat{x}\right)-\left(\tau_x^H+\tau_x^F\right) \hat{x}, \tag{A1}$$

where remember that the equilibrium $\hat{x}$ satisfies

$$\frac{1}{2}(1+\tau_1^H) y'\left(\hat{x}\right)=1+\frac{1}{2}\tau_x^H+\frac{1}{2}\tau_x^F. \tag{A2}$$

Our first result is that *regardless* of the equilibrium values of τ_1^H, τ_x^H, and τ_x^F, the surplus in equation (A1) will always be non-negative. To see this, note that using (A2) we can write

$$\left(1+\tau_1^H\right) y\left(\hat{x}\right) \geq (1+\tau_1^H)\hat{x} y'\left(\hat{x}\right)=2\hat{x}+\hat{x}\tau_x^H+\hat{x}\tau_x^F \geq \hat{x}\tau_x^H+\hat{x}\tau_x^F,$$

where we have used that the concavity of $y\left(\cdot\right)$ implies $y'\left(\hat{x}\right)\hat{x}<y\left(\hat{x}\right)$. Hence, the non-negativity constraint on the surplus can be ignored hereafter. Intuitively, no matter how distortionary trade taxes are, the level of investment x will adjust to ensure a positive joint surplus of the relationship.

Matters are not as simple with regards to the outside option of each producer. In particular, we are now careful to define this outside option as follows:

$$\max\left\{\frac{1}{2}\delta(1+\tau_1^H) y\left(x\right)-\frac{1}{2}\left(\tau_x^H+\tau_x^F\right) x, 0\right\}.$$

It is straightforward to see that whenever $\delta \to 0$, both producers will find it optimal to ignore the secondary market and simply throw away the amount x of input produced. In such a case, both outside options are zero and the (ex-post) payoff to both the producer and the supplier is given by

$$\pi^F=\pi^S=\frac{1}{2}\left((1+\tau_1^H) y\left(x\right)-\left(\tau_x^H+\tau_x^F\right) x\right), \tag{A3}$$

which is the same expression as in the main text. This implies that the analysis remains unchanged even when the non-negativity constraint is taken into account.

Finally, consider the model in section 5.1 with general bargaining power α in the secondary market. In such a case, it will also be true that for a low enough δ, the outside option of either home producers or foreign suppliers will be non-negative, in which case the secondary market will remain inactive. As a result, *stage*-3 payoffs will again be given by (A3), independent of α. In sum, when non-negativity constraints bind the solution of the model with a general α collapses to the solution of our Benchmark Model, where $\alpha=1/2$.

A.2. Second-Order Conditions

In this Appendix we provide a discussion of the second-order conditions of the main tariff setting games developed in the main text.

Second-Best Policy Choices in the Benchmark Model

It is easily verified that the second order conditions associated with the first-order conditions in (7) are satisfied. Simply note that evaluated at the equilibrium, we have

$$\begin{aligned}
\frac{\partial^2 W^W}{\partial \left(\tau_1^H\right)^2} &= \frac{\partial D_1}{\partial p_1^H} + y''\left(\hat{x}\right)\left(\frac{\partial \hat{x}}{\partial \tau_1^H}\right)^2 < 0 \\
\frac{\partial^2 W^W}{\partial \left(\tau_x\right)^2} &= y''\left(\hat{x}\right)\left(\frac{\partial \hat{x}}{\partial \tau_x}\right)^2 < 0 \\
\frac{\partial^2 W^W}{\partial \tau_x \partial \tau_1^H} &= y''\left(\hat{x}\right)\frac{\partial \hat{x}}{\partial \tau_1^H}\frac{\partial \hat{x}}{\partial \tau_x} < 0
\end{aligned}$$

and thus $\left(\partial^2 W^W/\partial\left(\tau_1^H\right)^2\right)\left(\partial^2 W^W/\partial\left(\tau_x\right)^2\right) - \left(\partial^2 W^W/\partial \tau_x \partial \tau_1^H\right)^2 = \left(\partial D_1/\partial p_1^H\right) y''\left(\hat{x}\right)\left(\partial \hat{x}/\partial \tau_x\right)^2 > 0.$

Nash Equilibrium Policy Choices in the Benchmark Model

We now consider the second-order conditions of the Nash equilibrium. Using equation (14), that is $\frac{1}{2}\left(1+\hat{\tau}_1^H\right) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\hat{\tau}_x^H + \frac{1}{2}\hat{\tau}_x^F$, we can simplify the first-conditions to conditions to obtain:

$$\begin{aligned}
\frac{\partial W^H}{\partial \tau_1^H} &= 0 = \tau_1^H \frac{\partial D_1}{\partial p_1^H} - \frac{1}{2} y\left(\hat{x}\right) + \left[y'\left(\hat{x}\right) - 1 - \tau_x^F\right]\frac{\partial \hat{x}}{\partial \tau_1^H}, \\
\frac{\partial W^H}{\partial \tau_x^H} &= 0 = \frac{1}{2}\hat{x} + \left[y'\left(\hat{x}\right) - 1 - \tau_x^F\right]\frac{\partial \hat{x}}{\partial \tau_x^H}, \text{ and} \\
\frac{\partial W^F}{\partial \tau_x^F} &= 0 = \frac{1}{2}\hat{x} + \tau_x^F \frac{\partial \hat{x}}{\partial \tau_x^F}.
\end{aligned}$$

Consider first the second-order condition for the choice of τ_x^F, i.e., $\partial^2 W^H/\partial\left(\tau_x^F\right)^2 < 0$. Differentiating the last expression above with respect to τ_x^F, we have

$$\frac{\partial^2 W^F}{\partial \left(\tau_x^F\right)^2} = \frac{3}{2}\frac{\partial \hat{x}}{\partial \tau_x^F} + \tau_x^F \frac{\partial^2 \hat{x}}{\partial \left(\tau_x^F\right)^2}. \tag{A4}$$

But using the implicit function theorem on (14), we have

$$\frac{\partial \hat{x}}{\partial \tau_x^F} = \frac{1}{y''\left(\hat{x}\right)}, \tag{A5}$$

which implies

$$\frac{\partial^2 \hat{x}}{\partial \left(\tau_x^F\right)^2} = -\frac{1}{\left(y''\left(\hat{x}\right)\right)^2} y'''\left(\hat{x}\right)\frac{\partial \hat{x}}{\partial \tau_x^F}. \tag{A6}$$

Using these expressions as well as $\tau_x^F = -\frac{1}{2}\hat{x}/\left(\partial \hat{x}/\partial \tau_x^F\right)$, we can write (A4) as

$$\frac{\partial^2 W^F}{\partial \left(\tau_x^F\right)^2} = \frac{\partial \hat{x}}{\partial \tau_x^F}\left(3 + \frac{\hat{x} y'''\left(\hat{x}\right)}{y''\left(\hat{x}\right)}\right),$$

which is negative only if $3+\hat{x}y'''(\hat{x})/y''(\hat{x})>0$. As an example, assume that $y(x)=x^{\eta}/\eta$, with $\eta\in(0,1)$. In such case, we have $y''(x)=(\eta-1)x^{\eta-2}$ and $y'''(x)=(\eta-2)(\eta-1)x^{\eta-3}$, and hence $3+\hat{x}y'''(\hat{x})/y''(\hat{x})=1+\eta$, which is indeed positive.

We can use this same condition to show that the possibility of foreign retaliation ($\tau_x^F>0$) worsens the hold-up problem and thus reduces $\hat{x}$. To see this, we can plug equation (A5) – which also applies to $\partial\hat{x}/\partial\tau_x^H$ – in the second first-order condition to obtain

$$0=\frac{1}{2}\hat{x}+\left[y'(\hat{x})-1-\tau_x^F\right]\frac{1}{y''(\hat{x})}=0.$$

This equation pins down $\hat{x}$ as a function of τ_x^F. Straightforward manipulation indicates that this expression is increasing in $\hat{x}$ as long as $3+\hat{x}y'''(\hat{x})/y''(\hat{x})>0$, which is the condition we have been imposing above. Because this expression is also increasing in τ_x^F, we can conclude that $\hat{x}$ will be lower whenever $\tau_x^F>0$ than whenever $\tau_x^F=0$.

The fact that in the Nash equilibrium we have $\hat{\tau}_1^H\neq 0$ implies that the second-order conditions for the choice of τ_1^H and τ_x^H are quite cumbersome to characterize, as they will now also involve properties of the demand function. Throughout the paper, we simply assume that they are satisfied without providing the exact conditions needed.

We next, however, develop a particular case of our model where the second order conditions are easy to characterize and simple comparative statics can be obtained. In particular, we make the simplifying assumption that demand for the final-good is perfectly elastic, which implies that $\hat{\tau}_1^H=0$. Under this assumption note that it is sufficient to check that $\partial^2W^H/\partial\left(\tau_x^H\right)^2<0$, which requires that

$$\frac{\partial^2W^H}{\partial\left(\tau_x^H\right)^2}=\frac{1}{2}\frac{\partial\hat{x}}{\partial\tau_x^H}+y''(\hat{x})\left(\frac{\partial\hat{x}}{\partial\tau_x^H}\right)^2+\left[y'(\hat{x})-1-\tau_x^F\right]\frac{\partial^2\hat{x}}{\partial\left(\tau_x^H\right)^2}<0.$$

Imposing $\partial W^H/\partial\tau_x^H=0$ to eliminate τ_x^F and plugging equations (A5) and (A6) – which also apply for τ_x^H –, we can simplify the above expression to:

$$\frac{\partial^2W^H}{\partial\left(\tau_x^H\right)^2}=\frac{1}{2}\frac{\partial\hat{x}}{\partial\tau_x^H}\left(3+\frac{\hat{x}y'''(\hat{x})}{y''(\hat{x})}\right)<0,$$

This again requires $3+\hat{x}y'''(\hat{x})/y''(\hat{x})>0$, which the same condition as in the choice of τ_x^F.

A.3. Location of the Secondary Market

We consider here the possibility that the secondary market for the foreign supplier is located in the foreign country. This implies that, in the event of disagreement with the final-good producer in H, the input supplier in F sells the inputs locally in the foreign country rather than exporting to an alternative buyer in H. There are a number of reasons to think that this possibility could be reflected in a richer model (e.g., as a result of search frictions associated with finding international partners on short notice that can be avoided with domestic matches), but rather than attempting to model these reasons explicitly we simply assume outright that there exists a secondary market in the foreign country (only) where a match with a local producer results in the production of an amount $y(x)$ of the generic good, and we return to the assumption of symmetric bargaining power.

The key difference relative to the Benchmark Model is in the outside options. The home producers now obtain no income in the secondary market, while foreign producers now obtain $\frac{1}{2}\delta(1+\tau_1^F)y(x)$ in that market, where τ_1^F is a foreign trade tax on the final good. Arguing as before, it is easy to see that the final-good producer in H now has a *stage*-3 payoff of

$$\frac{1}{2}\left[(1+\tau_1^H)-\frac{1}{2}\delta(1+\tau_1^F)\right]y(x)-\frac{1}{2}(\tau_x^H+\tau_x^F)x,$$

with the supplier in F now receiving a *stage*-3 payoff of

$$\frac{1}{2}\left[(1+\tau_1^H)+\frac{1}{2}\delta(1+\tau_1^F)\right]y(x)-\frac{1}{2}(\tau_x^H+\tau_x^F)x,$$

so that the *stage*-2 choice of $\hat{x}$ is now defined by

$$\frac{1}{2}\left[(1+\tau_1^H)+\frac{1}{2}\delta(1+\tau_1^F)\right]y'(\hat{x})=1+\frac{1}{2}(\tau_x^H+\tau_x^F), \tag{A6}$$

and hence the *stage*-1 payoffs of the home and foreign firm are given by

$$\begin{aligned}\pi^H &= \frac{1}{2}\left[(1+\tau_1^H)-\frac{1}{2}\delta(1+\tau_1^F)\right]y(\hat{x})-\frac{1}{2}(\tau_x^H+\tau_x^F)\hat{x}, \text{ and}\\ \pi^F &= \frac{1}{2}\left[(1+\tau_1^H)+\frac{1}{2}\delta(1+\tau_1^F)\right]y(\hat{x})-\frac{1}{2}(\tau_x^H+\tau_x^F)\hat{x}-\hat{x}.\end{aligned}$$

Anticipating that F may now have reason to alter p_1^F with its choice of τ_1^F (for reasons analogous to H's incentive to alter p_1^H with its choice of τ_1^H) and hence affect foreign consumer surplus $CS(p_1^F)$, and noting that none (or to be precise, a measure 0) of good 1 is actually produced in F in equilibrium, home and foreign welfare are then given by

$$\begin{aligned}W^H &= CS(p_1^H)+\pi^H+\tau_1^H[D_1(p_1^H)-y(\hat{x})]+\tau_x^H\hat{x}, \text{ and}\\ W^F &= CS(p_1^F)+\pi^F+\tau_1^F D(p_1^F)+\tau_x^F\hat{x}.\end{aligned}$$

The first-order conditions that define the Nash policies $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, $\hat{\tau}_1^{FN}$ and $\hat{\tau}_x^{FN}$ can be manipulated to yield

$$\begin{aligned}\hat{\tau}_1^{HN} &= \frac{\frac{1}{2}\hat{x}[y'(\hat{x})-\frac{y(\hat{x})}{\hat{x}}]}{-\partial D_1/\partial p_1^H},\\ \hat{\tau}_x^{HN} &= -\left[(1-\tau_1^H)-\frac{1}{2}\delta(1+\tau_1^F)\right]y'(\hat{x})-\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H}+\hat{\tau}_x^{FN},\\ \hat{\tau}_1^{FN} &= \frac{\frac{1}{4}\delta\hat{x}[y'(\hat{x})-\frac{y(\hat{x})}{\hat{x}}]}{-\partial D_1/\partial p_1^F}, \text{ and}\\ \hat{\tau}_x^{FN} &= -\frac{1}{2}\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F}.\end{aligned}$$

Again, the expression for $\hat{\tau}_1^{HN}$ is negative, and is similar to the expression derived in the Becnhmark Model and contained in (10). The intuition is also analogous to that in the Becnhmark Model: the home government finds it optimal to set a negative $\hat{\tau}_1^{HN}$ as a means of shifting surplus from foreign suppliers to the home country. The dual role that $\hat{\tau}_x^{HN}$ plays in alleviating the hold-up problem and at same time transferring surplus implies again that its sign is in general ambiguous.

This extension of the model delivers more interesting implications for the Nash policies adopted by the foreign government. First, as in the model the Becnhmark model, the foreign government has an incentive to set a positive export tax on the intermediate input ($\hat{\tau}_x^{FN} > 0$), because the foreign input supplier can pass part of this cost on to home producers by threatening not to deliver the intermediate input. The key for this is that the outside option for the supplier is not reduced one to one with τ_x^F. In the present variant of the model, this is not only due to less-than-full bargaining power for suppliers but also to the fact that the secondary market does not involve trade flows.

Second, and contrary to all of the models explored above, foreign taxes on the final good 1 can now affect the distribution of surplus between home and foreign producers. As a result, the foreign government now chooses to optimally balance the relative roles of $\hat{\tau}_x^{FN}$ and $\hat{\tau}_1^{FN}$ in extracting surplus from home firms in the same way that the home government balances $\hat{\tau}_x^{HN}$ and $\hat{\tau}_1^{HN}$ in extracting surplus from foreign firms. For the foreign government this implies the use of a foreign import tariff or export subsidy ($\hat{\tau}_1^{FN} > 0$) on the final good in order to raise p_1^F and thus improve the outside option (and bargaining position) of foreign suppliers.

Although we have shown that the location of the secondary market has implications for the Nash equilibrium values of home and foreign trade policies, it is important again to emphasize the two general features of our model that continue to hold in this extension as well. First, manipulating the above first-order conditions and applying the implicit function theorem to (A6), we find

$$y'(\hat{x}) = 1 - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H} > 1,$$

which indicates that again, under Nash policy choices, the international hold-up problem persists and the volume of international input trade is inefficiently low as a consequence. Second, as we have indicated our model predicts the equilibrium use of taxes in the final good market and these distortions arise as a result of each country's attempts to extract bargaining surplus from firms abroad. Once again therefore, the purpose of a trade agreement remains to help governments better solve these two problems.

Appendix B: A Competitive Benchmark

For comparison, we now develop the competitive analogue of our model. We suppose that foreign inputs are competitively supplied according to the supply curve

$$x_S^F \equiv x_S^F(p_x^F),$$

In country H, the final good 1 is produced according to the concave production function $y(x)$, and the marginal cost of production of final good 1 is given by

$$mc_1^H = \frac{p_x^H}{y'(x)}.$$

Competitive supply of final good 1 in country H is then determined according to $p_1^H = mc_1^H$ or

$$p_1^H = \frac{p_x^H}{y'(x_D^H)},$$

which implicitly defines x_D^H, the derived demand for the input x, as

$$x_D^H = y'^{-1}\left(p_x^H/p_1^H\right) \equiv x_D^H(p_1^H, p_x^H).$$

The pricing relationships are (with p_x^* the international or world/untaxed price):

$$p_1^H = 1 + \tau_1^H \equiv p_1^H(\tau_1^H); \quad p_x^H = p_x^* + \tau_x^H \equiv p_x^H(\tau_x^H, p_x^*); \quad p_x^F = p_x^* - \tau_x^F \equiv \; p_x^F(\tau_x^F, p_x^*).$$

The market-clearing condition in the world (home and foreign) x market is then given by $x_D^H = x_S^F$, or

$$x_D^H(p_1^H(\tau_1^H), p_x^H(\tau_x^H, p_x^*)) = x_S^F(p_x^F(\tau_x^F, p_x^*)), \tag{A7}$$

which determines $p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)$. Market-clearing input trade volume may then be written as $\hat{x}(p_1^H, p_x^H) \equiv x_D^H(p_1^H(\tau_1^H), p_x^H(\tau_x^H, p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)))$ or equivalently $\hat{x}(p_x^F) \equiv x_S^F(p_x^F(\tau_x^F, p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)))$. We also have $y(p_1^H, p_x^H) \equiv y(\hat{x}(p_1^H, p_x^H))$. Notice that (A7) can be totally differentiated to yield

$$\frac{\partial p_x^*}{\partial \tau_x^H} = \frac{-x_D^{H\prime}(p_1^H, p_x^H)}{x_D^{H\prime}(p_1^H, p_x^H) - p_1^H x_S^{F\prime}(p_x^F)} < 0; \quad \frac{\partial p_x^*}{\partial \tau_x^F} = \frac{-p_1^H x_S^{F\prime}(p_x^F)}{x_D^{H\prime}(p_1^H, p_x^H) - p_1^H x_S^{F\prime}(p_x^F)} > 0,$$

and so we have that

$$1 = \frac{\partial p_x^*}{\partial \tau_x^F} - \frac{\partial p_x^*}{\partial \tau_x^H}. \tag{A8}$$

The home welfare function may now be written as:

$$W^H = CS(p_1^H) + \gamma^H \int_0^{p_1^H} y(p, p_x^H)dp + (p_1^H - 1)[D_1^H(p_1^H) - y(p_1^H, p_x^H)] + (p_x^H - p_x^*)\hat{x}(p_1^H, p_x^H),$$

or

$$W^H \equiv W^H(p_1^H, p_x^H, p_x^*).$$

Similarly, the foreign welfare function may now be written as:

$$W^F = CS(1) + \gamma^F \int_0^{p_x^F} x_S^F(p)dp + (p_x^* - p_x^F)\hat{x}(p_x^F),$$

or

$$W^F \equiv W^F(p_x^F, p_x^*).$$

Using the fact that $W_{p_x^*}^F = -W_{p_x^*}^H = \hat{x}$, the efficiency frontier is defined by the three conditions:

$$\begin{aligned} W_{p_x^H}^H + [W_{p_x^H}^H + W_{p_x^F}^F]\frac{\partial p_x^*}{\partial \tau_x^H} &= 0, \\ -W_{p_x^F}^F + [W_{p_x^F}^F + W_{p_x^H}^H]\frac{\partial p_x^*}{\partial \tau_x^F} &= 0, \quad \text{and} \\ W_{p_1^H}^H + [W_{p_x^H}^H + W_{p_x^F}^F]\frac{\partial p_x^*}{\partial \tau_1^H} &= 0. \end{aligned}$$

Using (A8), it is easy to show that the first two first-order conditions are identical, and therefore determine the sum of τ_x^H and τ_x^F that is consistent with international efficiency.

To further interpret the conditions for efficiency, we multiply the first efficiency condition by $-[\frac{\partial p_x^*/\partial \tau_1^H}{\partial p_x^*/\partial \tau_x^H}]$ and add it to the third efficiency condition, so that we may then restate the two conditions for international efficiency as

$$\begin{aligned} W_{p_x^H}^H + [W_{p_x^H}^H + W_{p_x^F}^F]\frac{\partial p_x^*}{\partial \tau_x^H} &= 0, \text{ and} \\ W_{p_1^H}^H - W_{p_x^H}^H \cdot \frac{\partial p_x^*/\partial \tau_1^H}{\partial p_x^*/\partial \tau_x^H} &= 0. \end{aligned} \tag{A9}$$

The interpretation of (A9) is as follows. Let us begin with the second efficiency condition. On the left-hand side is the impact on home welfare of changes in the *mix* of τ_1^H and τ_x^H which hold fixed p_x^* – and hence, by (A7) and with τ_x^F and thus $p_x^F(\tau_x^F, p_x^*)$ unchanged, hold fixed as well the level of x_D^H and therefore the equilibrium level of input trade volume $\hat{x}$. Notice, though, that foreign welfare $W^F(p_x^F(\tau_x^F, p_x^*), p_x^*)$ is unaffected by such changes, because p_x^* is held fixed and τ_x^F is not changed and so, as already mentioned, $p_x^F(\tau_x^F, p_x^*)$ is held fixed as well. Hence, the second efficiency condition in (A9) says simply that, at internationally efficient choices of τ_1^H and τ_x^H, such changes can have no first-order effect on home welfare either. The first efficiency condition in (A9) then ensures that the sum of τ_x^H and τ_x^F achieves the efficient level of p_x^F, and hence the efficient level of input trade volume in light of the mix of τ_1^H and τ_x^H that the home country employs to deliver the chosen level of p_x^* and (with τ_x^F fixed) p_x^F.

Next consider the Nash policies. The associated first-order conditions are

$$\begin{aligned} W_{p_x^H}^H + [W_{p_x^H}^H + W_{p_x^*}^H]\frac{\partial p_x^*}{\partial \tau_x^H} &= 0, \\ -W_{p_x^F}^F + [W_{p_x^F}^F + W_{p_x^*}^F]\frac{\partial p_x^*}{\partial \tau_x^F} &= 0, \quad \text{and} \\ W_{p_1^H}^H + [W_{p_x^H}^H + W_{p_x^*}^H]\frac{\partial p_x^*}{\partial \tau_1^H} &= 0. \end{aligned} \tag{A10}$$

Using (A8) and $W^F_{p^*_x} = -W^H_{p^*_x}$, the first two Nash first-order conditions can be added together to yield:

$$W^H_{p^H_x} + [W^H_{p^H_x} + W^F_{p^F_x}]\frac{\partial p^*_x}{\partial \tau^H_x} + W^F_{p^*_x} = 0. \tag{A11}$$

Comparing (A11) to the first efficiency condition in (A9), the difference is the additional term $W^F_{p^*_x} > 0$ on the left-hand side of (A11), which implies that the sum $\tau^H_x + \tau^F_x$ is *inefficiently high* (the first-order condition for efficiency is negative at the Nash taxes), and therefore that the Nash level of input trade volume is inefficiently low in light of the mix of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x.

Next we multiply the initial first-order condition in (A10) by $-[\frac{\partial p^*_x/\partial \tau^H_1}{\partial p^*_x/\partial \tau^H_x}]$ and add it to the last first-order condition to get

$$W^H_{p^H_1} - W^H_{p^H_x} \cdot \frac{\partial p^*_x/\partial \tau^H_1}{\partial p^*_x/\partial \tau^H_x} = 0. \tag{A12}$$

Comparing (A12) to the second efficiency condition in (A9), we may conclude that the *mix* of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver its chosen level of p^*_x and hence p^F_x – and therefore by (A7), x^H_D and hence $\hat{x}$ – is internationally *efficient* (see Bagwell and Staiger, 2001, for an analogous observation).

Therefore, we may conclude that the single inefficiency in the Nash equilibrium in our competitive benchmark model is that the sum $\tau^H_x + \tau^F_x$ is *inefficiently high*, and hence that there is too little equilibrium input trade volume/input "market access": in the competitive benchmark model, the task of a trade agreement is thus to expand and secure market access to internationally efficient levels (see Bagwell and Staiger, 2001, 2002, for an interpretation of analogous findings from a market access perspective).

Next consider the political optimum conditions. Specifically, consider the hypothetical situation that governments are *not motivated* by the impact of their tariff choices on p^*_x, in the specific sense that $W^H_{p^*_x}\frac{\partial p^*_x}{\partial \tau^H_1} = W^H_{p^*_x}\frac{\partial p^*_x}{\partial \tau^H_x} \equiv 0$ and similarly for W^F. We then identify the tariffs that would be chosen unilaterally (i.e., non-cooperatively) by governments with these hypothetical preferences and ask whether these tariffs are efficient with respect to the actual government preferences. This is Bagwell and Staiger's (1999) original definition, and it is direct to show using (A10) that in our competitive benchmark model the following conditions define the political optimum:

$$W^H_{p^H_x} = 0, \; W^F_{p^F_x} = 0, \;\text{ and }\; W^H_{p^H_1} = 0. \tag{A13}$$

Clearly, as an examination of (A9) indicates, the political optimum defined in (A13) is efficient in this setting, so we now have shown that the standard terms-of-trade theory applies in a competitive-supplier version of our set-up. Moreover, notice from (A10) that an alternative definition of the political optimum, in which each government completely *ignores* the impact of its tariffs on p^*_x when it makes its unilateral tariff choices (in the particular sense that the home government acts as if $\frac{\partial p^*_x}{\partial \tau^H_1} = \frac{\partial p^*_x}{\partial \tau^H_x} \equiv 0$ and similarly the foreign government acts as if $\frac{\partial p^*_x}{\partial \tau^F_x} \equiv 0$) would lead to the same politically optimal tariffs as defined in (A13), and hence to the same conclusion. So it is the relationship-specific aspects of our supplier-producer setup – and the associated incomplete contracting and bilateral bargaining over price – that is crucial for the novel results that we derive in section 4.

References

Aghion, Philippe, Pol Antràs, and Elhanan Helpman (2007), "Negotiating Free Trade," *Journal of International Economics*, Vol. 73, No. 1, pp. 1-30.

Antràs, P. (2003), "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, 118:4, pp. 1375-1418.

Antràs, Pol (2005), "Incomplete Contracts and the Product Cycle," *American Economic Review*, 95:4, pp. 1054-1073.

Antràs, Pol, Luis Garicano and Esteban Rossi-Hansberg (2006), "Offshoring in a Knowledge Economy," *Quarterly Journal of Economics*, 121:1, pp. 31-77.

Antràs, Pol, and Elhanan Helpman (2004), "Global Sourcing," *Journal of Political Economy* 112:3, pp. 552-580.

Bagwell, K. and R.W. Staiger (1999), "An Economic Theory of GATT," *American Economic Review*, Vol. 89, No. 1, pp. 215-248.

Bagwell, K. and R.W. Staiger (2001), "Domestic Policies, National Sovereignty and International Economic Institutions," *Quarterly Journal of Economics*, Vol. 116, No. 2, pp. 519-562.

Bagwell, K. and R.W. Staiger (2002), *The Economics of the World Trading System*, MIT Press.

Bagwell, K. and R.W. Staiger (2006), "What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization," NBER Working Paper No. 12727.

Baron, D. (1994), "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review*, 88(1), pp. 33-47.

Bhagwati, J. and Ramaswami, V. K. (1963), "Domestic Distortions, Tariffs, and the Theory of Optimum Subsidy," *Journal of Political Economy*, 71:1, pp., 44-50.

Blanchard, E. (2007), "Foreign Direct Investment, Endogenous Tariffs, and Preferential Trade Agreements," working paper University of Virginia.

Bolton, P. and M. Dewatripont (2005), *Contract Theory*, MIT Press.

Bond, E. W. and C. Syropoulos (1996), "The Size of Trading Blocs: Market Power and World Welfare Effects," *Journal of International Economics*, 40:3, pp. 411-438.

Campa, Jose M., and Goldberg, Linda S. "The Evolving External Orientation of Manufacturing Industries: Evidence from Four Countries." *Federal Reserve Bank of New York Economic Policy Review* 3 (July 1997): 53-81.

Coase, Ronald H. (1937), "The Nature of the Firm," *Economica*, 4:16, pp. 386-405.

Deardorff, Alan V. (2001), "Fragmentation in Simple Trade Models," *North American Journal of International Economics and Finance*, 12(2), p. 121-137.

Diez, Federico (2006), "Tariffs on Final and Intermediate Goods under Global Sourcing," mimeo, University of Wisconsin.

Economist, The. "The Ins and Outs of Outing." (August 31, 1991): 54-56.

Ethier, W.J. (2004), "Political Externalities, Nondiscrimination, and a Multilateral World," *Review of International Economics*, Vol. 12, No. 3, pp. 303-320.

Feenstra, Robert C. "Integration of Trade and Disintegration of Production in the Global Economy." *Journal of Economic Perspectives* 12 (Autumn 1998), : 31-50.

Feenstra, Robert C., and Gordon H. Hanson, "Foreign Investment, Outsourcing and Relative Wages," in Robert C. Feenstra, Gene M. Grossman, and Douglas A. Irwin, eds., *Political Economy of Trade Policy: Essays in Honor of Jagdish Bhagwati* (Cambridge. MA: MIT Press, 1996a), pp. 89-127.

Feenstra, Robert C., and Hanson, Gordon H. "Globalization, Outsourcing, and Wage Inequality." *American Economic Review* 86 (May 1996b): 240-245.

Feenstra, Robert C., and Hanson, Gordon H. (2005), "Ownership and Control in Outsourcing to China: Estimating the Property Rights Theory of the Firm," *Quarterly Journal of Economics,* 120(2): 729-762.

Grossman, G.M. and Helpman, E. (1994), "Protection for Sale," *American Economic Review*, Vol. 84, No. 4, pp. 833-850.

Grossman, G.M. and E. Helpman (1996), "Electoral Competition and Special Interest Politics," The Review of Economic Studies, vol:63 pg:265.

Grossman, G.M. and Helpman, E. (2002), "Integration vs. Outsourcing in Industry Equilibrium," *Quarterly Journal of Economics* 117 (1), 85-120.

Grossman, G.M. and Helpman, E. (2005), "Outsourcing in a Global Economy," *Review of Economic Studies*, 72:1, pp. 135-159.

Grossman, Gene M. and Esteban Rossi-Hansberg (2006), "Trading Tasks: A Simple Theory of Offshoring," mimeo Princeton University.

Grossman, S. J., and O. D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94:4, pp. 691-719.

Hart, Oliver and John Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98:6, pp. 1119-1158.

Hummels, David; Ishii, Jun; and Yi, Kei-Mu. "The Nature and Growth of Vertical Specialization in World Trade." *Journal of International Economics*, 54 (June 2001): 75-96.

Johnson , H. G. (1953-54), "Optimum Tariffs and Retaliation," *Review of Economic Studies*, 21:2, pp. 142-153.

Johnson, H. G. (1965), "Optimal Intervention in the Presence of Domestic Distortions," in R. E. Caves, P. B. Kenen, and H.G. Johnson, eds., *Trade, Growth and the Balance of Payments: Essays in Honor of Gottfried Haberler*, Amsterdam: North-Holland, pp. 3-34.

Jones, Ronald W. Jones (2000), *Globalization and the Theory of Input Trade*, Cambridge: MIT Press.

Kennan, J. and Riezman, R. (1988), "Do Big Countries Win Tariff Wars?," *International Economic Review*, Vol. 29, No. 1, pp. 81-85.

Levchenko, Andrei (2007), "Institutional Quality and International Trade," *Review of Economic Studies*, 74:3, 791-819.

McLaren, J. (1997), "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade," *American Economic Review*, Vol. 87, No. 3, pp. 400-420.

Nunn, Nathan (2007), "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade," *Quarterly Journal of Economics*, Vol. 122, No. 2, pp. 569-600.

Ornelas, E. (forthcoming), "Feasible Multilateralism and the Effects of Regionalism," *Journal of International Economics*.

Ornelas, E. and J. Turner (forthcoming), "Trade Liberalization, Outsourcing, and the Hold Up Problem," *Journal of International Economics*.

Ramanarayanan, Ananth (2006), "International Trade Dynamics with Intermediate Inputs," mimeo University of Minnesota.

Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*, Free Press: New York.

Yarbrough, B.V. and R.M. Yarbrough (1992), *Cooperation and Governance in International Trade: The Strategic Organizational Approach*, Princeton University Press.

Yeats, Alexander J. "Just How Big Is Global Production Sharing?" In Arndt, Sven W., and Kierzkowski, Henryk eds., *Fragmentation: New Production Patterns in the World Economy*, Oxford University Press, 2001.