

# External Trade Diversion, Exclusion Incentives and the Nature of Preferential Trade Agreements

Paul Missios\*, Kamal Saggi<sup>†</sup> and Halis Murat Yildiz<sup>‡</sup>

## Abstract

In a game of endogenous trade agreements between three countries, we show that while the pursuit of customs unions (CUs) prevents global free trade from emerging as a coalition-proof Nash equilibrium, the pursuit of free trade agreements (FTAs) does not. This result reflects the relatively *flexible* nature of FTAs: whereas each FTA member can independently undertake further trade liberalization with respect to the non-member, CU members must do so as a group due to their common external tariff. By diverting members' exports away from the non-member, both types of trade agreements induce the non-member to voluntarily lower its import tariffs.

*Keywords:* Free Trade Agreement, Customs Union, Hub and Spoke Agreements, Free Trade, Optimal Tariffs. *JEL Classifications:* F11, F12.

---

\*Department of Economics, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. Phone: 416-979-5000 (ext 6186); e-mail: pmissios@ryerson.ca.

<sup>†</sup>Department of Economics, Vanderbilt University, Nashville, TN 37235-1828. Phone: 615-322-3237; e-mail: kamal.saggi@vanderbilt.edu.

<sup>‡</sup>Department of Economics, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. Phone: 416-979-5000 (ext 6689); e-mail: hyildiz@ryerson.ca.

# 1 Introduction

Preferential trade agreements (PTAs) are more popular than ever before while multilateral trade liberalization seems to have come to a standstill. The ever increasing popularity of PTAs can be gauged from the fact that as of 7 April 2015, 612 notifications of PTAs (counting goods and services separately) had been received by the GATT/WTO. Of these notifications, 406 PTAs are already in force with others scheduled for implementation in the near future. By contrast, despite fourteen years of fitful negotiations, the Doha Round of multilateral trade negotiations has failed to yield a bargain that is acceptable to all sides. Economists and policy-makers have long suspected that the contrasting fortunes of these two types of trade liberalization may be inter-related. More specifically, there is widespread concern that the formation of PTAs may undermine multilateral liberalization. Our objective in this paper is to investigate this issue with a finer lens by comparing the implications of the two most popular types of PTAs – i.e. free trade agreements (FTAs) and customs unions (CUs) – for the prospects of global free trade.<sup>1</sup>

As is well known, the central difference between an FTA and a CU is that members of a CU impose common external tariffs on non-members whereas FTA members adopt individually optimal tariffs. This difference in tariff setting behavior between the two types of PTAs has two important consequences. First, while an FTA member is free to enter into additional trade agreements with non-member countries without requiring consent from its existing FTA partners, a CU member can only do so if all other members also agree to participate in the new agreements.<sup>2</sup> In other words, FTA members enjoy more *flexibility* than CU members. Second, as noted in Bagwell and Staiger (1997a) and Bond and Syropoulos (1996), the coordination of tariffs within a CU allows members to *pool their market power*, thereby leading them to impose relatively higher external tariffs than FTAs.

The objective of this paper is to isolate the implications of these fundamental differences between FTAs and CUs for the prospects of free trade in the global economy. Our modeling approach follows that of Saggi and Yildiz (2010) and Saggi et al. (2013) under which PTAs emerge endogenously as the outcome of a game of trade liberalization between three countries that are free to pick their PTA partners as well as their tariff levels. We derive

---

<sup>1</sup>Roughly 90% of the existing PTAs take the form of FTAs, with CUs comprising the rest (Freund and Ornelas, 2010). However, the existing CUs do involve major trading areas of the world: the EU and much of Latin America (where MERCOSUR resides).

<sup>2</sup>For example, the United States has signed several FTAs since the ratification of NAFTA in 1995 that do not include Mexico and Canada as partners whereas the FTAs that the European Union has entered into over the years have required the approval and participation of all its existing members.

Nash equilibria of two games of PTA formation (called the FTA game and the CU game) and then isolate their set of coalition proof Nash equilibria (CPNE). Thus, our focus is on Nash agreements that are immune to self-enforcing coalitional deviations. As in Bernheim et al. (1987), all countries are free to arrange mutually beneficial deviations from Nash agreements and an initial coalitional deviation of two countries from a Nash equilibrium is self-enforcing if, taking the strategy of the third country as fixed, neither member of the deviating coalition has an incentive to further deviate to make itself better off.

In the first stage of either game, each country simultaneously announces the names of countries with whom it wants to sign a PTA. Next, given the trade policy regime, countries choose their optimal tariffs. In the final stage of the game, production, consumption and international trade take place. The mapping between first stage announcements and trade agreements is as follows. If no announcements match or the only matching announcement is in favor of no agreement, then no PTA is formed. A bilateral PTA arises if two countries announce each others' names and there is no other matching announcement. Since FTA members choose their external tariffs independently, two independent FTAs (or a hub and spoke trading regime) can arise in the FTA game and this happens if (i) a country (call it the hub) announces in favor of signing an FTA with the other two countries and (ii) the other two countries (spokes) announce in favor of signing an FTA with the hub country and at least one of them does not announce in favor of the other. By contrast, since CU members impose common external tariffs, a hub and spoke type of trading arrangement is *infeasible* in the CU game. In either game, free trade emerges only if each country seeks a PTA with both its trading partners.

Our paper addresses the important issue of when and why countries prefer to liberalize preferentially as opposed to doing so multilaterally. By contrast, existing literature generally takes one of two approaches: (i) it either takes PTAs to be exogenously given and compares incentives for multilateral liberalization in their presence to those in their absence or (ii) it simply asks whether a pair of countries benefit from entering into a PTA given that there exists no trade agreement between them, a comparison that does not tell us much about when and why they might deliberately choose to *exclude* others from their mutual trade liberalization. To address the exclusion incentive in a convincing manner, one needs a model that gives all countries an active voice in negotiations so that one can determine whether some countries prefer to exclude others from their mutual trade liberalization even though they wish to be included. We provide such a model and use it to assess the strength of the exclusion incentive under our two games as well as the ability of

member countries to exercise it in equilibrium.

The fact that external trade liberalization by a CU member is conditional on the approval of other members implies that, relative to a CU, an FTA is more susceptible to opportunistic unilateral deviations by member countries. To see this clearly, consider the FTA game and suppose countries  $i$  and  $j$  announce in favor of an FTA with each other. Then, country  $i$  is free to seek a *separate* FTA with country  $k$  in order to create a pair of independent bilateral FTAs (i.e. a hub and spoke arrangement) with itself as the hub and countries  $j$  and  $k$  as the spokes. By contrast, in the CU game, if countries  $i$  and  $j$  announce in favor of a CU with each other, the only way one of them can undertake further trade liberalization with country  $k$  is if they *both agree* to reduce their common external tariff. We show that this crucial difference between a CU and an FTA has important consequences for multilateral trade liberalization. In particular, in our three country model of symmetric countries, while free trade obtains as the unique CPNE of the FTA game, a CU between only two countries emerges as the unique CPNE of the CU game. Thus, in our model, the pursuit of CUs undermines global free trade whereas that of FTAs does not even though the exclusion incentive exists under both types of PTAs in the sense that the welfare of members of both types of PTAs is strictly higher than that under free trade.

The intuition behind this key result is as follows. Suppose each country announces in favor of forming an FTA with both other countries so that the resulting outcome is free trade. Due to the existence of an exclusion incentive in our model, two of these countries (say  $i$  and  $j$ ), taking the announcement of their complement as fixed (i.e. country  $k$  still announces in favor of FTAs with both  $i$  and  $j$ ), benefit if they jointly deviate to announcements wherein they call for an FTA with only each other. As per the concept of a CPNE, for this joint deviation to be self enforcing, a proper subset of the initially deviating countries (i.e.  $i$  or  $j$ ) should not have an incentive to deviate to another announcement, taking the announcement of the complement (country  $k$ ) as fixed. However, in the FTA game, each of the initially deviating country ( $i$  or  $j$ ) indeed has an incentive to deviate to an announcement in favor of FTAs with both countries since the welfare of a hub country exceeds that of the member of a single FTA. Therefore, the initial joint deviation of countries  $i$  and  $j$  from free trade is not self-enforcing and the CPNE of the FTA game yields free trade. By contrast, since members of a CU impose a common external tariff, two independent CUs (a hub and spoke type arrangement) are infeasible and the initial joint deviation of two countries that converts free trade to a bilateral CU is self-enforcing so that the CPNE of the CU game yields a bilateral CU. Thus, whereas the exclusion incentive

is reflected in the equilibrium of the CU game, it goes unexpressed in the FTA game due to the lure of a hub and spoke arrangement and the flexibility that FTA members have to pursue such an arrangement.

Our underlying trade model is an adapted version of the two-country comparative advantage based framework of Horn et al. (2010) where each country exports a unique good to the other two. This *competing importers* framework delivers a novel type of tariff complementarity: when two countries form a PTA (either an FTA or a CU), the excluded country voluntarily *reduces* its tariffs on them.<sup>3</sup> As a result, PTA members benefit not only from their mutual trade liberalization but also from the unilateral liberalization induced in the non-member country. This finding is reminiscent of the reciprocated unilateralism result of Krishna and Mitra (2005) who showed that unilateral tariff liberalization by a (large) country can result in a reciprocal tariff reduction by its smaller trading partner.<sup>4</sup> In their model, unilateral liberalization by the large country encourages the formation of an export lobby in the small country which then competes effectively with the import-competing lobby to lower tariffs and export taxes.

Our competing importers framework highlights a type of trade diversion that has generally been overlooked in the literature. Traditionally, trade diversion is defined as the increase in trade between PTA members that occurs at the expense of exports of non-members to the PTA, i.e., the traditional notion of trade diversion refers to the reduction in the volume of *imports* that PTA members source from non-members. By contrast, in our model, the formation of a PTA reduces the volume of *exports* of member countries to the non-member – a phenomenon we refer to as *external trade diversion*. Indeed, such trade diversion is precisely what makes it optimal for the non-member to reduce its tariffs on PTA members.<sup>5</sup> To the best of our knowledge, our model is the first to capture the effects of such external trade diversion on the import tariffs of the non-member country and the associated welfare implications of such a change, issues that cannot be addressed by models that take the tariffs of non-members to be exogenously given or assume a trading structure under which PTA formation does not affect them.<sup>6</sup>

---

<sup>3</sup>By contrast, existing literature has tended to focus on how the formation of a PTA can induce member countries to reduce their external tariffs. See Bagwell and Staiger (1997a, 1997b and 1998), Bond et al. (2004), and Estevadeordal et al. (2008).

<sup>4</sup>See also Coates and Ludema (2001) for a theory of trade policy leadership based on repeated interaction between a large and a small country. In their model, a large country can undertake unilateral liberalization in order to induce reciprocal trade liberalization by the small country.

<sup>5</sup>In our model, this logic would also apply to unilateral liberalization by any country so long as it is preferential in nature (i.e. extended only to one country).

<sup>6</sup>See, for example, Bagwell and Staiger (1997a, 1997b, 1998), Krishna (1998), Ornelas (2005a and

To isolate the market power effect of a CU from their relative lack of flexibility, in section 4.1 we also consider a scenario where CU members are not allowed to raise their tariffs above pre-existing levels. This experiment is well motivated on policy grounds: Article XXIV of GATT – the key clause that sanctions PTAs in the WTO – forbids member countries of a PTA from raising tariffs on non-members. When CUs are constrained in this manner, we find that a bilateral CU continues to arise in the CPNE of the CU game thereby thwarting global free trade. This finding implies that while the restriction on external tariffs mandated by Article XXIV softens the negative impact of a CU on outsiders, it does not eliminate the exclusion incentive that gives rise to a CU in the first place. Thus, it is the relative flexibility of FTAs over CUs that helps the prospects of global free trade and not their weaker market power.

While the present paper follows the analytical approach of Saggi and Yildiz (2010) and Saggi et al. (2013), it highlights an important conceptual difference between the two types of PTAs that has not been explored in these papers (or in the rest of the literature on PTAs). More specifically, the conceptual point that the relative inflexibility of a CU makes it coalitionally more stable than an FTA has never really been explicitly proven in the literature since most existing papers generally focus on only one type of a PTA.

Unlike the present paper, both Saggi and Yildiz (2010) and Saggi et al. (2013) analyze PTA formation in a *competing exporters* model of trade in which each country imports the same good from two other countries. Both of these papers primarily focus on the issue of whether PTAs are building or stumbling blocs for free trade by comparing outcomes of two games – one where discriminatory PTAs are allowed and another where only non-discriminatory multilateral agreements are permitted. This is not an issue that we address here. Saggi and Yildiz (2010) focus on FTAs and find that in their competing exporters framework, no two countries have an incentive to exclude the third country from their mutual trade liberalization. In their model, under symmetry the outcome is global free trade whereas under asymmetry the outcome is an agreement between the two larger countries not because they wish to exclude the third country but rather because the third country *chooses* to stay out on its own accord. In other words, the exclusion incentive never arises in Saggi and Yildiz (2010). By contrast, when the model in the present paper is extended to an asymmetric setting, even in the FTA game the two larger countries deliberately choose to exclude the smaller country even though it would be better off being included in the

---

2005b), and Saggi and Yildiz (2010). In general, the literature has tended to focus on the reduction of market access experienced by non-members due to the formation of a PTA.

agreement. A major real world concern regarding FTAs has always been whether they can become devices for exclusion. By allowing for endogenous agreements between all players, our paper demonstrates that this concern regarding FTAs and exclusion is a real one.

The CU game analyzed by Saggi et al. (2013) in a competing exporters model of trade delivers free trade under symmetry. By contrast, in the present paper, the outcome of the CU game even under symmetry is a bilateral CU. Furthermore, the present paper shows that the exclusion incentive of a CU exists even when its members are not allowed to fully exercise their market power by raising their tariffs above preexisting levels. This result shows that it is the relative lack of flexibility of CUs relative to FTAs that prevents the obtainment of global free trade in the CU game, as opposed to their higher market power. Finally, note that though the exclusion incentive plays a role in the CU game of Saggi et al. (2013) when countries are asymmetric, the magnitude of this incentive is quite different in the present paper: in our competing importers framework, the reduction in the external tariff of the non-member induced by external trade diversion reinforces the exclusion incentive whereas no such effect arises in a competing exporters framework.

The rest of the paper is organized as follows. Section 4.1 examines the CU game under the restriction that CU members cannot raise their tariff above preexisting levels. In section 4.2, we consider a game where countries can choose between the two types of PTAs and show that the unique CPNE of this game delivers a bilateral CU. In section 4.3, we show that this basic insight holds even when countries are not fully symmetric. Section 5 contains concluding remarks.

## 2 Tariffs and trade

Our underlying economic framework is an adapted version of the two-country model of Horn et al. (2010). We consider a perfectly competitive world with three large countries:  $z = i, j$ , and  $k$  and three (non-numeraire) goods:  $g = I, J$ , and  $K$  and a numeraire good  $v_0$ . On the demand side, the representative citizen's utility function is linear in the numeraire good and separable in the non-numeraire goods:

$$U(\mathbf{v}, v_0) = u(\mathbf{v}) + v_0, \tag{1}$$

where  $\mathbf{v} = [v_I, v_J, v_K]$  is the consumption vector for the three non-numeraire goods,  $v_0$  denotes the consumption of the numeraire good, and  $u(\mathbf{v})$  is quadratic and additively separable in the three non-numeraire goods. The demand for good  $g$  in country  $z$  is then

given by

$$d_z^g(p_z^g) = \alpha - p_z^g \quad (2)$$

where  $p_z^g$  denotes the consumer price of good  $g$  in country  $z$ . Assuming that the population in each country is a continuum of measure one, we can write the consumer surplus associated with good  $g$  in country  $z$  as:

$$CS_z^g(p_z^g) = u_z^g[d_z^g(p_z^g)] - p_z^g d_z^g(p_z^g) \quad (3)$$

On the supply side, as in Horn et al. (2010), labour ( $l$ ) is the only factor of production which is employed in the production of the numeraire good that is produced one-for-one from labor. The supply of labor is assumed to be large enough that the numeraire good is always produced in a positive amount; therefore the equilibrium wage is equal to one.

Each non-numeraire good is produced from labor with diminishing returns. In particular, we assume the following production function for non-numeraire good  $g$  in country  $z$ :  $Q_z^g = \sqrt{2\lambda_z^g l_g}$ , where  $Q_z^g$  is the production of good  $g$  in country  $z$  and  $l_g$  is the labor employed in the production of good  $g$ . The supply function of good  $g$  in country  $z$  is as follows:

$$s_z^g(q_z^g) = \lambda_z^g q_z^g \quad (4)$$

where  $q_z^g$  denotes the producer price for good  $g$  in country  $z$ .

We assume that there exists a symmetric comparative advantage structure across countries:  $\lambda_i^I = \lambda_j^J = \lambda_k^K = 1 + \lambda$  while  $\lambda_i^J = \lambda_i^K = \lambda_j^I = \lambda_j^K = \lambda_k^I = \lambda_k^J = 1$ . In other words, each country has a comparative advantage in one good while having a comparative disadvantage in the other two: each country exports the good that is indexed by the same uppercase letter as the identity of the country. For example, country  $i$  exports good  $I$  while importing good  $J$  from country  $j$  and good  $K$  from country  $k$ . Thus, there are two *competing importers* for each non-numeraire good and the model is Ricardian in nature with diminishing returns in the production of each good. Country  $z$ 's producer surplus in good  $g$  as follows:

$$PS_z^g(q_z^g) = \int s_z^g(q_z^g) dq_z^g = \frac{1}{2} \lambda_z^g (q_z^g)^2 \quad (5)$$

As a representative scenario for all goods and countries, consider good  $I$  (i.e. the good in which country  $i$  has a comparative advantage). Let  $t_{ji}$  be the tariff imposed by country  $j$  on its imports of good  $I$  from country  $i$ .<sup>7</sup> Given that all countries are large, the world

<sup>7</sup>We assume that tariff revenues for each good are redistributed uniformly to all individuals.

price of good  $I$  depends on the tariffs chosen by countries  $j$  and  $k$  but to simplify notation we suppress the dependence of prices on tariffs and simply denote the price of good  $I$  by  $p_i^I$ .

Due to the absence of any tariff in country  $i$  on good  $I$ , the consumer and producer prices of good  $I$  in country  $i$  are equal:  $q_i^I = p_i^I$ . As there is no domestic taxation for the import competing sectors, producer and consumer prices are also equal:  $q_i^g = p_i^g$ , where  $g \neq I$ . Finally, ruling out prohibitive tariffs yields the following no-arbitrage conditions for good  $I$ :

$$p_j^I = p_i^I + t_{ji} \text{ and } p_k^I = p_i^I + t_{ki} \quad (6)$$

Let  $m_j^I$  and  $m_k^I$  be the imports of good  $I$  by countries  $j$  and  $k$ :

$$m_j^I(p_j^I) = d(p_j^I) - s_j^I(p_j^I) \text{ and } m_k^I(p_k^I) = d(p_k^I) - s_k^I(p_k^I) \quad (7)$$

Similarly, let  $x_j^I$  and  $x_k^I$  denote country  $i$ 's exports (of good  $I$ ) to countries  $j$  and  $k$  where

$$x_j^I = s_i^I(p_i^I) - d(p_i^I) - m_k^I(p_k^I) \quad (8)$$

and

$$x_k^I = s_i^I(q_i^I) - d(p_i^I) - m_j^I(p_j^I) \quad (9)$$

Market clearing for good  $I$  requires that country  $i$ 's exports to a country equals the imports of that country:

$$x_z^I = m_z^I \text{ where } z = j, k \quad (10)$$

Before proceeding with the derivation of optimal tariffs, it is useful to highlight some important features of the model. Since each country exports a unique good in the model, a country's tariff on one of its trading partners has no impact on the volume of its imports from its second trading partner. This implies that a country's external tariffs are independent of one another (since they apply to different goods, each with its own demand function).<sup>8</sup> Second, if two countries liberalize trade only towards one another, they import more from each other and start exporting less to the third country – a phenomenon which we call *external trade diversion*. As we will see below, this reduction in the volume of exports to the third country in turn has implications for its optimal tariffs.

---

<sup>8</sup>One consequence of this feature is that the MFN principle of non-discrimination plays no role in our model.

## 2.1 Optimal Nash tariffs

Country  $z$ 's welfare is defined as the sum of consumer surplus, producer surplus, and tariff revenue over all goods:

$$w_z = \sum_g CS_z^g(p_z^g) + \sum_g PS_z^g(q_z^g) + \sum_{h \neq z} t_{zh} m_z^H \quad (11)$$

where  $H$  corresponds to the good exported by country  $h$ .

In the absence of any trade agreement, each country chooses its tariffs to maximize its welfare. To derive optimal tariffs, we follow the approach of Feenstra (2004) and Broda et al. (2008).

Consider country  $k$ 's tariff problem. Differentiating  $w_k$  with respect to  $t_{ki}$ , we obtain:

$$\frac{\partial w_k}{\partial t_{ki}} = t_{ki} \frac{\partial m_k^I}{\partial p_k^I} \frac{\partial p_k^I}{\partial t_{ki}} - m_k^I \frac{\partial p_i^I}{\partial t_{ki}} \quad (12)$$

The first term of the above first order condition is the efficiency cost of the tariff (i.e. the marginal deadweight loss from the tariff) while the second term is the terms of trade effect, that is, the reduction in the price of good  $i$  that accrues to country  $i$  ( $p_i^I$ ) multiplied by the quantity of country  $k$ 's imports from country  $i$ .

The optimal ad-valorem tariff is computed where (12) equals zero:

$$\frac{\partial w_k}{\partial t_{ki}} = 0 \Rightarrow \frac{t_{ki}}{p_i^I} = \frac{\frac{\partial p_i^I}{\partial t_{ki}} \frac{m_k^I}{p_i^I}}{\frac{\partial m_k^I}{\partial p_k^I} \frac{\partial p_k^I}{\partial t_{ki}}} \quad (13)$$

The above expression can be interpreted in two different ways. Note first that since  $m_k^I = x_k^I$  we must have

$$\frac{\partial m_k^I}{\partial p_k^I} \frac{\partial p_k^I}{\partial t_{ki}} = \frac{\partial x_k^I}{\partial t_{ki}}$$

Substituting this into (13) shows that country  $k$ 's optimal ad-valorem tariff equals the inverse of the elasticity of the export supply curve of country  $i$  to country  $k$ , denoted by  $\varepsilon_{ik}$ :

$$\frac{t_{ki}}{p_i^I} = \frac{1}{\varepsilon_{ik}} = \left[ \frac{\partial x_k^I}{\partial p_i^I} \frac{p_i^I}{x_i^I} \right]^{-1} \quad (14)$$

For an alternative interpretation of the optimal tariff, we can rearrange (13) and write

the optimal tariff formula for country  $k$  as follows:

$$\frac{t_{ki}}{p_i^I} = \frac{1}{\mu_k^I} \frac{\partial p_i^I}{\partial t_{ki}} \left[ \frac{\partial p_k^I}{\partial t_{ki}} \right]^{-1} \quad \text{where } \mu_k^I = \frac{\partial m_k^I}{\partial p_k^I} \frac{p_k^I}{m_k^I} \quad (15)$$

As can be seen from above, the optimal tariff is also equal to the inverse of the elasticity of import demand of good  $I$  in  $k$  ( $\mu_k^I$ ), times the ratio of the change in the relative world price and domestic price of imports. Given that import demand elasticity  $\mu_k^I < 0$ , the fact that country  $k$ 's tariff on good  $I$  drives down the local price of the good in country  $i$  (i.e.  $\frac{\partial p_i^I}{\partial t_{ki}} < 0$ ) while raising it locally ( $\frac{\partial p_k^I}{\partial t_{ki}} > 0$ ), the optimal tariff imposed by country  $k$  is positive. Note that the term  $\frac{\partial p_i^I}{\partial t_{ki}}$  captures the *terms of trade gain* of the tariff since it informs us how country  $k$ 's tariff on good  $I$  affects the price collected by country  $i$  while  $\frac{\partial p_k^I}{\partial t_{ki}}$  refers to the *pass through* of the tariff since it tells us how the domestic price of good  $I$  in country  $k$  varies with its tariff on country  $i$ .<sup>9</sup>

Using the demand and supply functions in equations (2) and (4) as well as equations (6) through (10), the equilibrium prices of good  $I$  in country  $i$  and in importing country  $j$  equal:

$$p_i^I = \frac{3\alpha - 2 \sum_{z \neq i} t_{zi}}{\lambda + 6} \quad \text{and} \quad p_j^I = \frac{3\alpha - 2t_{ki} + (4 + \lambda)t_{ji}}{\lambda + 6} \quad (16)$$

As is clear from equation (16), the price of good  $I$  in country  $i$  decreases in the degree of comparative advantage  $\lambda$  (supply effect) and the tariffs it faces in export markets (terms of trade effect). Similarly, the prices of good  $I$  in country  $j$  increases with its own tariff whereas it decreases with the tariff of the rival importer (i.e. country  $k$ ). Using the above price equations, we can explicitly calculate the terms of trade gain and the pass through of import tariffs.<sup>10</sup> We have:

$$\frac{\partial p_i^I}{\partial t_{ji}} = \frac{\partial p_i^I}{\partial t_{ki}} < 0 \quad (17)$$

i.e. the tariffs imposed by countries  $j$  and  $k$  lower the price collected by country  $i$ 's exporters and

$$\frac{\partial p_j^I}{\partial t_{ji}} = \frac{\partial p_k^I}{\partial t_{ki}} < 1 \quad (18)$$

i.e. the pass through from tariffs to local prices in importing countries is incomplete – i.e. the local price in a country does not increase one-to-one with its import tariff.

---

<sup>9</sup>Broda et al. (2008) provide evidence that the importers with market power indeed use it in setting their non-cooperative trade policy.

<sup>10</sup>Using the prices in (16), the volume of trade as a function of tariffs can be easily calculated.

The first order condition in (12) can be written as:

$$\frac{\partial w_k}{\partial t_{ki}} = \frac{2[\alpha\lambda - t_{ki}(\lambda + 4)(\lambda + 8) + 4t_{ji}]}{(\lambda + 6)^2} \quad (19)$$

It is immediate from the above first order condition that we have positively sloped reaction functions, i.e., tariffs imposed by different countries on the same good (i.e. their common import) are *strategic complements* in our model:

$$\frac{\partial t_{ki}}{\partial t_{ji}} > 0 \quad (20)$$

The intuition for why tariffs of different countries end up being strategic complements in our model is easy to see: an increase in the tariff country  $j$  imposes on country  $i$  increases the volume of country  $i$ 's exports to country  $k$  thereby increasing the latter's ability to manipulate its terms of trade. Simultaneous solution of first order conditions for countries  $j$  and  $k$  leads to the following optimal Nash tariffs (which are equal due to symmetry):

$$t_{ji}^* = t_{ki}^* = t^* = \frac{\alpha\lambda}{\lambda^2 + 12\lambda + 28} \quad (21)$$

## 2.2 How PTA formation affects tariffs

If countries  $i$  and  $j$  form an FTA  $\langle ij \rangle$ , they remove their tariffs on each other ( $t_{ij} = t_{ji} = 0$ ) and impose their individually optimal external tariffs on the non-member country by solving:

$$\left. \frac{\partial w_i}{\partial t_{ik}} \right|_{t_{ij}=0} = 0 \text{ and } \left. \frac{\partial w_j}{\partial t_{jk}} \right|_{t_{ji}=0} = 0 \quad (22)$$

Since member countries  $i$  and  $j$  are competing importers of good  $K$ , the elimination of their internal tariffs due to the FTA has *no impact* on their tariffs on imports of good  $K$  and thus the optimal external tariff of FTA members, denoted by  $t(ij)$ , is the same as that under no agreement, denoted by  $t^*$ . We have  $t(ij) = t^*$ .<sup>11</sup>

On the other hand, since countries compete over imports, the formation of an FTA between two countries changes the first order conditions of the tariff choice problem of the

---

<sup>11</sup>Thus, the model does not exhibit the type of tariff complementarity described in Bagwell and Staiger (1997a, 1997b) and some other models of PTAs.

non-member country. We have:

$$\left. \frac{\partial w_k}{\partial t_{ki}} \right|_{t_{ji}=0} = \frac{2[\alpha\lambda - t_{ki}(\lambda + 4)(\lambda + 8)]}{(\lambda + 6)^2} \quad (23)$$

Since we know from (20) that tariffs on the same good imposed by different importers are strategic complements, the reduction of  $t_{ij}$  and  $t_{ji}$  to zero induces country  $k$  to lower its external tariff on the imports from country  $i$  and country  $j$ . We have:

$$t_k(ij) = \frac{\alpha\lambda}{(\lambda + 4)(\lambda + 8)} < t^*$$

We can now state:

**Proposition 1A:** *The external trade diversion caused by a bilateral FTA induces the non-member to lower its tariffs on members (i.e.  $t_k(ij) < t^*$ ) while it has no effect on the external tariffs of members (i.e.  $t(ij) = t^*$ ).*

The general intuition behind the above proposition is as follows. Since the removal of internal tariffs under the FTA  $\langle ij \rangle$  leads to an increase in the imports of country  $j$  from country  $i$ , the export supply curve of country  $i$  to country  $k$  shifts to the left. As a result, the equilibrium world price of good  $I$  rises while the equilibrium exports of country  $i$  to country  $k$  decline. Since export supply curves are linear in our model, the elasticity of export supply curves of countries  $i$  and  $j$  facing country  $k$  rises due to the formation of the bilateral FTA  $\langle ij \rangle$ :

$$\varepsilon_{ik}(ij) - \varepsilon_{ik}(\Phi) = 2/\lambda > 0 \quad (24)$$

which in turn implies that country  $k$ 's optimal tariff on imports from countries  $i$  and  $j$  under the FTA  $\langle ij \rangle$  is lower than its optimal Nash tariff  $t^*$ .

For an alternative interpretation of Proposition 1A, consider the optimal tariff formula in (15). Note that the FTA  $\langle ij \rangle$  leads to an upward movement along the import demand curves of country  $k$  for goods  $i$  and  $j$  leading to a higher elasticity of import demand (common due to symmetry). Thus, the removal of the internal tariffs under the FTA  $\langle ij \rangle$  leads to a lower terms of trade gain and a higher tariff pass through in country  $k$ :

$$\left. \frac{\partial p_i^I}{\partial t_{ki}} \right|_{t_{ji}=0} > \left. \frac{\partial p_i^I}{\partial t_{ki}} \right|_{t_{ji}=t^*} \quad \text{and} \quad \left. \frac{\partial p_k^I}{\partial t_{ki}} \right|_{t_{ji}=0} > \left. \frac{\partial p_k^I}{\partial t_{ki}} \right|_{t_{ji}=t^*}$$

both of which tend to lower its optimal tariff. Analogous reasoning explains why the two spoke countries end up imposing lower tariffs on each other relative to the status quo:

under the hub and spoke arrangement, each spoke exports more to the hub and less to the other spoke relative to the status quo, which in turn lowers the ability of both spokes to manipulate their terms of trade vis-à-vis one another.

Next, we consider the formation of a CU between countries  $i$  and  $j$ , denoted by  $\langle ij_u \rangle$ . Like FTA members, CU members remove tariffs on each other. However, unlike FTA members, CU members impose a jointly optimal external tariff on the non-member. Under the CU  $\langle ij_u \rangle$  members solve:

$$\max_t w_i(ij^u) + w_j(ij^u) \text{ subject to } t_{ij} = t_{ji} = 0 \quad (25)$$

As the common market of a CU is larger than that of its members, the combined market power of the CU exceeds that of the individual member countries. Since the terms-of-trade externalities across members are internalized by a CU and tariffs on the same good across countries are complementary in our model, the optimum external tariffs of members rise following the formation of CU.<sup>12</sup>

To gain further insight, let  $m^K$  be the total import demand of good  $K$  in countries  $i$  and  $j$ :

$$m^K = \sum_{z=i,j} [d(p_z^K) - s_z^K(q_z^K)] \quad (26)$$

while  $x^K$  denote the exports of good  $K$  (from country  $k$ ) to the common market of countries  $i$  and  $j$ :<sup>13</sup>

$$x^K = s_k^K(q_k^K) - d(p_k^K) \quad (27)$$

The equilibrium world price of good  $K$  is determined by the market clearing condition:

$$x^K = m^K \quad (28)$$

Using the market clearing prices and quantities, the first order condition for the welfare maximization problem in (25) can be written as:

$$\frac{\partial[w_i(ij^u) + w_j(ij^u)]}{\partial t} = \frac{4[2\alpha\lambda - t(\lambda + 2)(\lambda + 10)]}{(\lambda + 6)^2} = 0 \quad (29)$$

---

<sup>12</sup>Olarreaga et al. (1999) provide evidence that the terms-of-trade externalities among Mercosur's members were internalized in its external tariffs.

<sup>13</sup>Note that with a CU we are back to two country set-up of Feenstra (1994) and Broda et al. (2008).

Solving this yields the optimal tariff of the CU  $\langle ij_u \rangle$ :<sup>14</sup>

$$t(ij^u) = \frac{2\alpha\lambda}{\lambda^2 + 12\lambda + 20} \quad (30)$$

Since both countries import the same good from the non-member country, the market power effect of a CU emphasized by Bagwell and Staiger (1997a) arises here. A CU allows members to pool their market power and extract a larger terms of trade gain from the non-member leading to an increase in their tariffs:

$$t(ij^u) > t(ij) = t^* \quad (31)$$

It is also immediate that since member countries reduce their internal tariffs to zero, like an FTA, a bilateral CU also induces the non-member to lower its tariffs on members (and exactly to the same level):<sup>15</sup>

$$t_k(ij^u) = t_k(ij) < t^* \quad (32)$$

It is straightforward to establish that the external trade diversion caused by a CU has qualitative similar effects as an FTA:

**Proposition 1B:** *The formation of a CU induces members to raise their external tariffs (i.e.  $t(ij^u) > t^*$ ) and the non-member to lower them:  $t_k(ij^u) < t^*$ .*

To see the general picture behind the above result more clearly, suppose that we start with the status quo where all countries impose the optimal tariff  $t^*$  on each other. It is obvious that the export supply of country  $k$  to the CU's common market lies to the right of its individual export supply curves to the markets of countries  $i$  and  $j$  since it applies to a common larger market. Similarly, following the formation of the CU  $\langle ij_u \rangle$ , the common import demand of member countries is larger relative to individual import demands. Therefore, market clearing under the CU  $\langle ij_u \rangle$  occurs at a lower  $p_k^K/x_k^K$  ratio relative to the status quo. Since export supply curves are linear, the elasticity of country  $k$ 's export supply curve falls due to the formation of the CU  $\langle ij_u \rangle$ . As a result, the optimal common external tariff under CU  $\langle ij_u \rangle$  is higher than the individually optimal tariffs of countries  $i$  and  $j$ .

It is worth noting that Article XXIV of GATT requires PTA members to not raise their external tariffs on non-member countries. For now, we ignore this tariff restriction imposed

---

<sup>14</sup>It is straightforward to confirm that the optimal ad-valorem tariff of the CU equals the inverse of the elasticity of country  $k$ 's export supply curve:  $t/p_k^K = 1/\varepsilon_k(ij_u)$ .

<sup>15</sup>Thus, Proposition 1 holds regardless of whether the PTA is an FTA or a CU.

by Article XXIV and assume that a CU can impose its optimal tariffs. In section 4.1, we examine the implications of this tariff restriction on welfare and on the prospects of global free trade.

### 3 Endogenous preferential trade agreements

We now describe our three stage game of trade liberalization. In the first stage, each country simultaneously announces the names of countries with whom it wants to sign a PTA. This stage determines the underlying trade policy regime. Next, given the trade policy regime, countries choose their optimal tariffs. Finally, consumption, production and international trade take place. As noted before, a key difference between the two types of PTAs is the relative flexibility of FTAs: while an FTA member is free to sign another FTA with a non-member without needing consent of an existing FTA partner, a CU member cannot do so since all CU members must have a common external tariff. To capture the implications of this important difference between an FTA and a CU, we now derive and compare the equilibrium outcomes of the two games.

#### 3.1 Free trade agreements

In the first stage of the FTA game, each country simultaneously announces the names of countries with whom it wants to sign an FTA. Country  $i$ 's announcement is denoted by  $\sigma_i$  and its strategy set  $S_i$  consists of four possible announcements:

$$S_i = \{\{\phi, \phi\}, \{j, \phi\}, \{\phi, k\}, \{j, k\}\}$$

where  $\{\phi, \phi\}$  denotes an announcement in favor of no FTA with either trading partners,  $\{j, \phi\}$  in favor of an FTA with only country  $j$ ;  $\{\phi, k\}$  in favor of an FTA with only country  $k$ ; and  $\{j, k\}$  in favor of FTAs with both of them.

It is important to note that we employ a game of announcements or proposals. In our game, a country does not announce in favor of a specific trade agreement but rather names partners with whom it wants to form such agreements. Furthermore, since a trade agreement requires consent from both sides, we posit the following mapping between various announcements profiles and the types of trade agreements that countries can form:

(i) No two announcements match or the only matching announcements are  $\{\phi, \phi\}$ . All of these announcement profiles lead to no agreement  $\langle \Phi \rangle$  wherein all countries impose their

optimal Nash tariffs on one another.

(ii) Two countries announce each others' name and there is no other matching announcement: i.e., the following three conditions hold: (a)  $j \in \sigma_i$  and  $i \in \sigma_j$ ; (b)  $i \notin \sigma_k$  and/or  $k \notin \sigma_i$  and (c)  $j \notin \sigma_k$  and/or  $k \notin \sigma_j$ . All announcements satisfying these three conditions yield an FTA between countries  $i$  and  $j$  denoted by  $\langle ij \rangle$  under which members impose zero tariffs on each other and the optimal external tariff  $t^*$  on the non-member.

(iii) Country  $i$  announces in favor of signing an FTA with countries  $j$  and  $k$  while countries  $j$  and/or  $k$  announce only in favor of signing an FTA with country  $i$ : i.e.  $\sigma_i = \{j, k\}$ ;  $i \in \sigma_j$ ; and  $i \in \sigma_k$  while  $k \notin \sigma_j$  and/or  $j \notin \sigma_k$ . This set of announcements yields a pair of independent FTAs (i.e. a hub and spoke trading regime) with  $i$  as the common member denoted by  $\langle ij, ik \rangle$  (or simply  $\langle ih \rangle$ ) under which countries  $j$  and  $k$  impose the optimal external tariff on each other while imposing no tariffs on country  $i$  who also imposes no tariffs on either of them.

(iv) All countries announce each others' names. These announcements yield global free trade  $\langle F \rangle$  under which all countries eliminate their tariffs on each other.<sup>16</sup>

Note that since an FTA between two countries can arise only if it is mutually acceptable to both sides, multiple announcement profiles can map into the same agreement. For example, the FTA  $\langle ij \rangle$  obtains when (i) countries  $i$  and  $j$  call only each other, *regardless* of the nature of country  $k$ 's announcement: if  $\sigma_i = \{j, \phi\}$  and  $\sigma_j = \{i, \phi\}$ , then  $\langle ij \rangle$  obtains for all four possible announcements on the part of country  $k$ , i.e., for  $\sigma_k = \{\phi, \phi\}$ ,  $\{i, \phi\}$ ,  $\{\phi, j\}$  and  $\{i, j\}$  so that country  $k$ 's announcement has no bearing upon the outcome when neither of the other two countries' announces its name; (ii) countries  $i$  and  $j$  announce each other's name and either one or both of them also announce country  $k$  but country  $k$  does not reciprocate: i.e. all of the following types of announcements map into the FTA  $\langle ij \rangle$ : (a)  $\sigma_i = \{j, k\}$  and  $\sigma_j = \{i, \phi\}$  but  $i \notin \sigma_k$  or (b)  $\sigma_i = \{j, \phi\}$  and  $\sigma_j = \{i, k\}$  but  $j \notin \sigma_k$  or (c)  $\sigma_i = \{j, k\}$  and  $\sigma_j = \{i, k\}$  but  $\sigma_k = \{\phi, \phi\}$ .

### 3.1.1 Welfare effects of various trade agreements

Let country  $i$ 's welfare as a function of the underlying trade agreement  $a$  be denoted by  $w_i(a)$  and let  $\Delta w_i(a - b)$  denote the difference between country  $i$ 's welfare under trade agreements  $a$  and  $b$ :  $\Delta w_i(a - b) \equiv w_i(a) - w_i(b)$ . Also, let  $m$  denote a member country of the bilateral FTA  $\langle ij \rangle$  so that  $m = i$  or  $j$ .

---

<sup>16</sup>We should note here that while a scenario under which each country signs an FTA with both its trading partners is institutionally distinct from global free trade, from the perspective of applied tariffs the two are the same since all countries set their tariffs to zero under both scenarios.

We first state the following preliminary result that explains the welfare effects of various types of trade agreements:

**Lemma 1:** *Let  $m, \tilde{m} = i, j$  denote member countries of the FTA  $\langle ij \rangle$ . The welfare impact of various types of trade agreements on members and non-members is as follows:*

- (i)  $\Delta w_k(ij - \Phi) < 0 < \Delta w_m(ij - \Phi)$ .
- (ii)  $\Delta w_m(F - ij) < 0 < \Delta w_k(F - ij)$ .
- (iii)  $\Delta w_{\tilde{m}}(mh - ij) < 0 < \Delta w_m(mh - ij)$ .
- (iv)  $\Delta w_k(mh - ij) > 0$ .
- (v)  $\Delta w_m(F - mh) < 0 < \Delta w_{\tilde{m}}(F - mh)$ .

Part (i) says that, given that there exists no trade agreement, a pair of countries have an incentive to form a bilateral FTA and the formation of such an FTA makes the non-member country worse-off. In our model, the formation of a bilateral FTA does not affect the non-member's producer surplus in export markets since external tariffs of FTA members are the same as those under the status quo (i.e.  $t(ij) = t^*$ ). Furthermore, since the external trade diversion caused by an FTA lowers the non-member's ability to manipulate its terms of trade vis-à-vis member countries (which is reflected, in turn, in its reduced external tariff), the non-member is worse off relative to the status quo. Since aggregate world welfare increases due to the trade liberalization undertaken by FTA member countries, it follows that the formation of an FTA makes member countries better off at the expense of the non-member.

The result stated in part (ii) of Lemma 1 establishes the existence of an *exclusion incentive* since a pair of countries prefer a bilateral FTA to free trade. Furthermore, since world welfare under free trade is higher than that under a bilateral FTA, it follows that the non-member country is better off under free trade relative to a bilateral FTA. The forces that give rise to the exclusion incentive can be understood as follows. Relative to free trade, each member country of an FTA has the ability to manipulate its terms of trade vis-à-vis the non-member while also being able to free ride on the terms of trade effect of the non-member's tariff on its FTA partner. These benefits of exclusion are somewhat tempered by the fact that, relative to free trade, an FTA member faces positive tariffs in the non-member's market (but recall that these tariffs are lower than those under the status quo owing to the external trade diversion caused by the FTA). All in all, the two positive effects of exclusion on FTA members dominate the negative effect so that a pair of countries benefit if they can successfully exclude the third country from their bilateral trade agreement.

Part (iii) says the welfare of a hub country is higher than that of the member of a single bilateral FTA whereas that of a spoke country is lower. It is important to note that, starting from a single FTA, the external trade diversion caused by the hub country's second FTA reduces the volume of exports flowing from the original non-member (who becomes a spoke) to the original member (who becomes the other spoke). This trade diversion reduces the original non member's ability to manipulate its terms of trade which shows up as a reduction in its optimal external tariff.

Part (iv) simply says that it is better to be a spoke than to be a non-member facing a bilateral FTA. When the non-member country forms an FTA with one of the member countries, it not only obtains free access to the hub country's market but also faces lower tariffs in the other spoke country due to external trade diversion. However, it also gives free access to its market to the hub country while keeping its tariff on the other spoke country unchanged. Overall, the former positive effects dominates the latter negative effect.

Finally, part (v) says that while each spoke country is better off under free trade, the hub country is worse off. A comparison of hub country  $m$ 's welfare under  $\langle mh \rangle$  relative to  $\langle F \rangle$  yields the following: (i) the hub country's producer surplus is equal under the two regimes since its producers face zero tariffs under both regimes and are afforded no protection in the home market (ii) its domestic welfare is higher under  $\langle mh \rangle$  relative to  $\langle F \rangle$  since it benefits from the positive terms of trade effects of tariffs that the spokes impose on each other – indeed, it is able to import goods from both spokes at prices that are *lower* than those under free trade. As a result, the hub country is strictly better off relative to free trade. To see why the spokes are worse off relative to free trade, first note that aggregate global welfare is strictly higher under free trade. Given that the hub country is strictly better off relative to free trade and the fact that the welfare of the two spokes is equal (due to symmetry), both spokes are worse off relative to free trade.

### 3.1.2 Nash Equilibria of the FTA game

Using the welfare comparisons stated in Lemma 1, we can show that the Nash equilibria of the FTA game are as follows:

**Proposition 2A:** *The Nash equilibrium announcement profiles of the FTA game and the associated trade agreements are as follows:*

<i>Announcement Profiles</i>	<i>Trade Agreement</i>
$\Omega^\Phi \equiv \{\sigma_i = \{\phi, \phi\}, \sigma_j = \{\phi, \phi\}, \sigma_k = \{\phi, \phi\}\}$	$\langle \Phi \rangle$
$\Omega^{ij} \equiv \{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{\phi, \phi\}\}$	$\langle ij \rangle$
$\Omega^{ih} \equiv \{\sigma_i = \{j, k\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, \phi\}\}$	$\langle ih \rangle$
$\Omega^F \equiv \{\sigma_i = \{j, k\}, \sigma_j = \{i, k\}, \sigma_k = \{i, j\}\}$	$\langle F \rangle$

The logic behind Proposition 2A is as follows. It is straightforward that the announcement profile  $\Omega^\Phi$  is a Nash equilibrium since no country has an incentive to announce another's name if the latter does not announce its name in return. Next consider  $\Omega^{ij}$ . Note from part (i) of Lemma 1 that neither member country ( $i$  or  $j$ ) has an incentive to unilaterally change its announcement from that which it makes under  $\Omega^{ij}$  since doing so only leads to no agreement  $\langle \Phi \rangle$ , under which its welfare is lower. Similarly, given that neither country  $i$  nor country  $k$  announces its name, country  $k$  has no incentive to alter its announcement from  $\sigma_k = \{\phi, \phi\}$  since doing so has no bearing on the resulting trade agreement. Thus, the announcement profile  $\Omega^{ij}$  yielding a bilateral FTA is a Nash equilibrium.

Now consider the announcement profile  $\Omega^{ih}$  associated with the hub and spoke regime  $\langle ih \rangle$ . First note that the hub country  $i$  has no incentive to unilaterally change its announcement from  $\{j, k\}$  to  $\{j, \phi\}$  or  $\{\phi, k\}$  since doing so translates into a deviation from the hub and spoke regime  $\langle ih \rangle$  where  $i$  is the hub country to  $\langle ij \rangle$  or  $\langle ik \rangle$  respectively. Similarly, the hub country has no incentive to alter its announcement to  $\{\phi, \phi\}$  since doing so yields  $\langle \Phi \rangle$  under which it is also worse off. Now consider the incentive of a spoke country to deviate from  $\Omega^{ih}$ . Neither spoke country (say  $j$ ) has an incentive to unilaterally deviate from its announcement  $\{i, \phi\}$  to  $\{\phi, \phi\}$  or  $\{i, k\}$  since the former deviation translates into a deviation from  $\langle ih \rangle$  where  $j$  is a spoke country to  $\langle ik \rangle$  where  $j$  is a non-member country while the latter deviation does not alter the trading regime, i.e., it remains  $\langle ih \rangle$ . As a result, neither the hub nor the spokes have an incentive to unilaterally alter their announcements from  $\Omega^{ih}$  that yields the hub and spoke regime  $\langle ih \rangle$ .

Finally, consider the announcement profile  $\Omega^F$  that yields global free trade  $\langle F \rangle$ . Since world welfare is the highest under  $\langle F \rangle$ , each country is better off under  $\langle F \rangle$  to  $\langle \Phi \rangle$  and this preference together with part (ii) of Lemma 1 implies that a country (say  $k$ ) has no incentive to unilaterally deviate from its announcement  $\{i, j\}$  to  $\{\phi, \phi\}$  since doing so alters the trade regime from from  $\langle F \rangle$  to  $\langle ij \rangle$  under which it is worse off.

Moreover, we know from part (v) of Lemma 1 that a country (say  $k$ ) has no incentive to unilaterally change its announcement from  $\{i, j\}$  to  $\{i, \phi\}$  or  $\{\phi, j\}$  since this translates into a deviation from  $\langle F \rangle$  to  $\langle ih \rangle$  or  $\langle jh \rangle$  under both of which country  $k$  is a spoke country.

Thus,  $\Omega^F$  is also a Nash equilibrium.

Observe from Proposition 2A that there is a unique announcement profile that supports each agreement as a Nash equilibrium and that the profile itself is the most *parsimonious* one. For example, even though  $\{\sigma_i = \{\phi, j\}, \sigma_j = \{\phi, \phi\}, \sigma_k = \{\phi, \phi\}\}$  also maps to  $\langle \Phi \rangle$ , such an announcement profile does not constitute a Nash equilibrium. To see why, simply note that given these announcements, country  $j$  has an incentive to alter its announcement from  $\sigma_j = \{\phi, \phi\}$  to  $\sigma_j = \{i, \phi\}$  in order to form the bilateral FTA  $\langle ij \rangle$ . Similarly, it is worth considering briefly as to why  $\{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, \phi\}\}$  is not a Nash equilibrium profile even though, just as the announcement profile  $\Omega^{ij}$ , it also maps into the FTA  $\langle ij \rangle$ . Under this announcement profile, given the announcements of countries  $j$  and  $k$ , as per part (iii) of Lemma 1, country  $i$  has an incentive to alter its announcement to  $\sigma_i = \{j, k\}$  so as to obtain the trade agreement  $\langle ij \rangle$  under which it is the hub. Using analogous reasoning, we rule out all other non-parsimonious announcements as candidates for Nash equilibria in the Appendix.

It is worth noting that except for  $\Omega^\Phi$  – which is Pareto dominated by the announcement profile  $\Omega^F$  that supports global free trade – the *other three Nash equilibria are all Pareto efficient*. To see why, simply note from part (ii) of Lemma 1 that  $\Omega^F$  does not Pareto dominate  $\Omega^{ij}$  since FTA members prefer the FTA  $\langle ij \rangle$  to global free trade. Similarly, part (v) of Lemma 1 implies that  $\Omega^F$  also fails to Pareto dominate  $\Omega^{ih}$  since the hub country is worse off under free trade while the spokes are better off. Finally, to see why  $\Omega^{ih}$  and  $\Omega^{ij}$  cannot be Pareto ranked, observe from part (iii) of Lemma 1 that while country  $i$  is better off under  $\langle ih \rangle$  relative to  $\langle ij \rangle$ , country  $j$  is worse off.

### 3.1.3 Coalition Proof Nash Equilibria

In what follows, we refine the set of Nash equilibria by isolating those Nash equilibria that are coalition proof. Bernheim et al. (1987) state that “in an important class of “noncooperative” environments, it is natural to assume that players can freely discuss their strategies, but cannot make binding commitments. In such cases, any meaningful agreement between the players must be self-enforcing. Although the Nash best-response property is a necessary condition for self-enforceability, it is not sufficient - it is in general possible for coalitions to arrange plausible, mutually beneficial deviations from Nash agreements.” Allowing countries to discuss their strategies regarding which trade agreements they intend to form is eminently desirable in the present context since countries considering bilateral trade agreements certainly have the capacity to communicate with one another without

necessarily having the ability to make binding commitments regarding their future plans. Following Bernheim et al. (1987): “... an agreement is coalition-proof if and only if it is Pareto efficient within the class of self-enforcing agreements. In turn, an agreement is self-enforcing if and only if no proper subset (coalition) of players, taking the actions of its complement as fixed, can agree to deviate in a way that makes all of its members better off.” Therefore, a coalition proof Nash equilibrium (CPNE) is immune to all self-enforcing coalitional deviations.

We are now ready to state one of our key results (see the appendix for its proof):

**Proposition 3A:** *The only coalition proof Nash equilibrium of the FTA game is the announcement profile  $\Omega^F$  that yields global free trade.*

The key message of this result is that even though a pair of countries benefit from excluding the third country from their trade agreement, *they are unable to exercise this exclusion incentive in equilibrium.* The flexible nature of FTAs plays a crucial role in ensuring that the exclusion incentive goes unexercised in the FTA game. To see why, suppose the announcement profile is  $\Omega^F$  – i.e. each country announces in favor of an FTA with both its trading partners. Starting at  $\Omega^F$ , as part (ii) of Lemma 1 indicates, two countries (say  $i$  and  $j$ ) have an incentive to exclude the third country by jointly altering their announcements in way that the announcement profile changes from  $\Omega^F$  to  $\{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, j\}\}$  thereby altering the associated trade regime from free trade to the bilateral FTA  $\langle ij \rangle$ . However, from Lemma 1 we know that the most preferred trading arrangement from each country’s perspective is a hub and spoke regime under which it is the hub. It follows then that, holding constant the announcement of the excluded country at  $\sigma_k = \{i, j\}$ , each member of the deviating coalition has an incentive *to alter its announcement to form a separate FTA with the excluded country.* For example, country  $i$  has an incentive to alter its announcement from  $\sigma_i = \{j, \phi\}$  to  $\sigma_i = \{j, k\}$  which alters the announcement profile from  $\{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, j\}\}$  to  $\{\sigma_i = \{j, k\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, j\}\}$  and the associated trade regime from  $\langle ij \rangle$  to  $\langle ih \rangle$ . Since the welfare of a hub is higher than that of a member country of a single FTA – see part (iii) of Lemma 1 – the original coalitional deviation of countries  $i$  and  $j$  from is not self-enforcing. Thus, in a nutshell, since it is better to be a (i) hub under a hub and spoke regime than to be a member of a single FTA and (ii) a spoke than to be a non-member facing a bilateral FTA, in the FTA game no two countries are able to successfully exclude the third country by forming a bilateral FTA: the lure of a hub and spoke trading arrangement ends up delivering free trade as the only CPNE of the FTA game.

It is important to acknowledge that when evaluating whether a particular coalitional deviation is self-enforcing or not, the concept of a CPNE holds the announcement of the excluded country constant. Thus, when country  $i$  alters its announcement from  $\sigma_i = \{j, \phi\}$  to  $\sigma_i = \{j, k\}$  in order to secure the hub and spoke regime  $\langle ih \rangle$ , the CPNE concept rules out the possibility of a pact between countries  $j$  and  $k$  to enter into an FTA with each other. Thus, by imposing this type of a restriction, the concept of a CPNE captures a relatively limited type of self-enforceability. Is such a restriction plausible in our model of FTAs?<sup>17</sup> As Bernheim et al. (1987) note, one possible justification for this approach is as follows: “...informational problems may cripple any attempt to form coalitions consisting of both members and nonmembers from some other deviating coalition.” This justification is fairly reasonable in our set-up since discussions amongst a pair of countries regarding a potential FTA between them are likely to be private and an excluded country may lack verifiable information about the content of such discussions. Real world FTAs are relatively complex and multi-faceted and third parties may not have the means to uncover what exactly is being negotiated between potential member countries. The full details of an FTA generally become public knowledge only after it has been ratified by member countries and the actual signing of an FTA by two countries is rather different from early negotiations between them, detailed information about which may not be easily available to outsiders.

An alternative way of understanding Proposition 3A is as follows. In our three player game, to isolate the CPNE of the FTA game we only need to consider coalitional deviations from each Nash equilibrium to other Nash equilibria since, by definition, any coalitional deviations from a Nash equilibrium announcement profile to other profiles that are not Nash equilibria are susceptible to unilateral deviations on the part of at least one member of the deviating coalition and therefore do not need to be considered. A Nash equilibrium profile from which no coalition wishes to deviate to another Nash equilibrium profile would then be coalition proof. Using this logic, we can immediately rule out the announcement profile  $\Omega^\Phi$  as a CPNE since countries  $i$  and  $j$  benefit from a joint deviation to  $\Omega^{ij}$ , which is a Nash equilibrium. Next, note that  $\Omega^{ij}$  also cannot be coalition proof since parts (iii) and (iv) of Lemma 1 imply that countries  $i$  and  $k$  benefit by jointly deviating to the Nash equilibrium profile  $\Omega^{ih}$ . Furthermore,  $\Omega^{ih}$  is also not coalition proof since, as per part (v)

---

<sup>17</sup>We thank an anonymous referee for raising this point. While the concept of CPNE suffers from this limitation, alternative equilibrium concepts – such as Strong Nash – have problems of their own. As Bernheim et al. (1987) note, the concept of Strong Nash gives coalitions too much freedom in choosing their deviations since a Strong Nash equilibrium has to be immune to *all* coalitional deviations, even to those that are not self-enforcing in nature. As a result, Strong Nash equilibria rarely exist. Indeed, the set of Strong Nash equilibria of our FTA game is empty.

of Lemma 1, countries  $j$  and  $k$  benefit if they jointly deviate from  $\Omega^{ih}$  to  $\Omega^F$ , which is a Nash equilibrium. Finally,  $\Omega^F$  is indeed coalition proof because any joint deviation from  $\Omega^F$  to an alternative Nash announcement profile *makes at least one member of the deviating coalition worse off*. For example, as per part (ii) of Lemma 1, the deviation from  $\Omega^F$  to  $\Omega^{jj}$  makes country  $k$  worse off; that from  $\Omega^F$  to  $\Omega^{ih}$  makes countries  $j$  and  $k$  worse off; and that from  $\Omega^F$  to  $\Omega^\Phi$  makes all countries worse off. As a result,  $\Omega^F$  is immune to all coalitional deviations.

We now analyze the CU game.

### 3.2 The CU game

Suppose the PTA under consideration is a CU as opposed to an FTA. As under the FTA game, at the first stage of the CU formation game each country announces the names of countries with whom it wants to form a CU. Country  $i$ 's announcement is denoted by  $\sigma_i$  and its strategy set  $S_{iu}$  consists of four possible announcements:

$$S_{iu} = \{\{\phi, \phi\}, \{j_u, \phi\}, \{\phi, k_u\}, \{j_u, k_u\}\} \quad (33)$$

where  $\{\phi, \phi\}$  denotes an announcement in favor of no CU with either trading partners,  $\{j_u, \phi\}$  in favor of a CU with only country  $j$ ;  $\{\phi, k_u\}$  in favor of a CU with only country  $k$ ; and  $\{j_u, k_u\}$  in favor of a CU that includes both its trading partners.

The mapping between various announcements profiles and the CUs that can arise is as follows: (i) when no two announcements match or the only matching announcements are  $\{\phi, \phi\}$  we obtain no agreement  $\langle \Phi \rangle$ ; (ii) a CU between countries  $i$  and  $j$  denoted by  $\langle ij_u \rangle$  is formed if they announce each others' names and there is no other matching announcement, i.e.,  $\langle ij_u \rangle$  is formed if the following three conditions hold: (a)  $j_u \in \sigma_i$ ,  $i_u \in \sigma_j$  and (b)  $k_u \notin \sigma_i$  and/or  $i_u \notin \sigma_k$  and (c)  $k_u \notin \sigma_j$  and/or  $j_u \notin \sigma_k$ ; (iv) free trade  $\langle F \rangle$  obtains iff all countries announce each other's names. Recall that the equivalent of a hub and spoke trading regime cannot arise under the CU game due to the fact that CU members coordinate their external tariffs.

In order to derive Nash equilibria of the CU game, we first summarize the relevant welfare comparisons:

**Lemma 2:** *Let  $m = i, j$  denote a member country of the CU  $\langle ij_u \rangle$ . Then we have: (i)  $\Delta w_k(ij_u - \Phi) < 0 < \Delta w_m(ij_u - \Phi)$  and (ii)  $\Delta w_m(ij_u - F) > 0$ .*

Lemma 2 can be interpreted much like Lemma 1. Part (i) simply says that the formation

of a CU makes members better off and the non-member worse off. In fact, since CU members *pool their market power* and extract a larger terms of trade gain from the non-member, the welfare of a CU member is higher than that of an FTA member while that of the non-member is lower:

$$\Delta w_k(ij_u - ij) < 0 < \Delta w_m(ij_u - ij)$$

Part (ii) of Lemma 2 says that the exclusion incentive is very much at play in the CU game. In fact, *the exclusion incentive of a CU is stronger than that of an FTA*, which is another implication of the higher market power of a CU:

$$0 < \Delta w_m(ij - F) < \Delta w_m(ij_u - F)$$

### 3.2.1 Equilibria of the CU game

We can now state:

**Proposition 2B:** *The Nash equilibrium announcement profiles of the CU game and the associated trade agreements are as follows:*

- (i)  $\Omega_u^\Phi \equiv \{\sigma_i = \{\phi, \phi\}, \sigma_j = \{\phi, \phi\}, \sigma_k = \{\phi, \phi\}\}$  which yields no agreement  $\langle \Phi \rangle$ .
- (ii) All announcement profiles satisfying  $j_u \in \sigma_i, i_u \in \sigma_j$  and any one of the following three conditions yield the bilateral CU  $\langle ij_u \rangle$ : (1)  $k_u \notin \sigma_i$  and  $k_u \notin \sigma_j$  or (2)  $k_u \notin \sigma_i$  and  $j_u \notin \sigma_k$  or (3)  $k_u \notin \sigma_j$  and  $i_u \notin \sigma_k$ .
- (iii)  $\Omega_u^F \equiv \{\sigma_i = \{j_u, k_u\}, \sigma_j = \{i_u, k_u\}, \sigma_k = \{i_u, j_u\}\}$  which yields global free trade  $\langle F \rangle$ .

The logic underlying Proposition 2B follows directly from Lemma 2. The announcement profile  $\Omega_u^F$  is a Nash equilibrium since no country has an incentive to unilaterally alter its announcement because doing so results in free trade being replaced by a bilateral CU between the other two countries under which it is worse off. Furthermore,  $\Omega_u^\Phi$  is a Nash equilibrium since no country has an incentive to announce another's name given that the latter does not reciprocate. The key difference relative to the FTA game is described in part (ii) of Proposition 2B. Note that as long as countries  $i$  and  $j$  announce each other's names but not that of country  $k$ , the resulting agreement is the bilateral CU  $\langle ij_u \rangle$  regardless of what country  $k$  announces. When countries  $i$  and  $j$  announce each other's names and one of them announces country  $k$  but country  $k$  does not make any matching announcement, the resulting agreement is again  $\langle ij_u \rangle$ . As such, a variety of different announcement profiles are consistent with a bilateral CU. Following Lemma 2,  $\langle ij_u \rangle$  is the most preferred trade regime

of member countries. Therefore, neither country  $i$  nor  $j$  has an incentive to unilaterally change its announcement and a unilateral change in country  $k$ 's announcement does not alter the existing trade agreement  $\langle ij_u \rangle$  since a CU member cannot have an independent trade agreement with the third country due to the requirement of a common external tariff.

Following our discussion of the FTA game, it is easy to see that  $\Omega_u^\Phi$  is not a CPNE of the CU game. To see why, simply note that starting at  $\Omega_u^\Phi$ , countries  $i$  and  $j$  benefit by altering their announcements to  $\sigma_i = \{j, \phi\}$  and  $\sigma_j = \{i, \phi\}$  which converts no agreement  $\langle \Phi \rangle$  to the bilateral CU  $\langle ij_u \rangle$  under which both of them are better off (see Lemma 2). Since the announcement profile  $\{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{\phi, \phi\}\}$  is a Nash equilibrium of the CU game, it is immune to unilateral deviations on the part of countries  $i$  and  $j$ .

We now argue that  $\Omega_u^F$  is also *not* a CPNE of the CU game. To see why, simply consider the coalitional deviation of countries  $i$  and  $j$  from  $\Omega_u^F$  to  $\Omega_u^1 \equiv \{\sigma_i = \{j, \phi\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, j\}\}$ . Observe from Proposition 2B that this coalitional deviation alters the trade regime from free trade to  $\langle ij_u \rangle$  and since CU members enjoy higher welfare than that under free trade, neither member country has an incentive to unilaterally alter its announcements from that which it makes under  $\Omega_u^1$ . The key difference relative to the FTA game is that the joint deviation of countries  $i$  and  $j$  from each of them announcing both their trading partners names' to calling only each other is now self-enforcing since, holding constant country  $k$ 's announcement, neither of them has an incentive to further deviate unilaterally. Thus,  $\Omega_u^F$  fails to be a CPNE of the CU game.

Thus, the only remaining candidate for a CPNE is the CU  $\langle ij_u \rangle$ . To see why it must be a CPNE, simply note that, starting with any Nash equilibrium announcement profile that yields  $\langle ij_u \rangle$ , member countries have no incentive to jointly alter their announcements to either obtain  $\langle \Phi \rangle$  or  $\langle F \rangle$  since they are worse off under either of these outcomes. Furthermore, since at least one of the members does not announce the non-member's name in any announcement profile that supports  $\langle ij_u \rangle$ , no coalition of a member and a non-member can obtain free trade by jointly altering their announcements. Thus, we can state the following:

**Proposition 3B:** *Only those announcement profiles that yield a bilateral CU constitute the coalition proof Nash equilibria of the CU game.*

The difference between results reported in Propositions 3A and 3B is driven by the relatively flexible nature of FTAs compared to CUs. In the FTA game, if two countries ( $i$  and  $j$ ) try to jointly exclude the third country by forming a bilateral FTA then each of them has an incentive to sign an independent FTA with the excluded country thereby making itself a hub. The ability to act on this incentive acts as a deterrent for the other

initially deviating country (say  $j$ ) since it is worse off as a spoke under  $\langle ih \rangle$  relative to free trade and thus the initial joint deviation from the announcement profile  $\Omega^F$  to  $\Omega^{ij}$  does not occur. However, unlike the FTA game, no such deterrent exists under the CU game since a CU member cannot form an independent agreement with the excluded country without the consent of its CU partner. Therefore, our model suggests that the pursuit of PTAs is compatible with the goal of achieving global free trade only when PTAs take the form of FTAs.

We should note here that our non-cooperative approach rules out transfers between countries. It is worth discussing briefly the case where the use of such transfers is permitted. Since joint welfare is maximized under global free trade, if international transfers could be freely used then the outcome of the CU game would also be free trade. In the real world, it is probably the case that countries can use transfers but they also face some constraints regarding how large such transfers can be. If this is the case, our analysis shows that while the FTA game would yield free trade, the CU game will fail to do so if the benefit of the exclusion to two countries exceeds the maximum level of transfer that the non-member can provide to induce CU members to move to global free trade – i.e. free trade would not arise if the maximum level of the transfer that an excluded country can provide to each CU member to induce it to move to free trade falls below  $\Delta w_m(ij_u - F)$ .

## 4 Further analysis and discussion

In what follows, motivated by Article XXIV of GATT, we first examine the consequences of restricting CU members from raising their external tariff above the pre-existing level  $t^*$ .<sup>18</sup> Next, holding this restriction in place, we analyze two different extensions. In the first extension, we analyze a more general game where countries can choose between FTAs and CUs. In the second extension, we consider a scenario where demand is asymmetric across countries in order to examine how allowing for such type of asymmetry alters our key results.

---

<sup>18</sup>Part 5(a) of Article XXIV of GATT states that “with respect to a customs union, or an interim agreement leading to a formation of a customs union, the duties and other regulations of commerce imposed at the institution of any such union or interim agreement in respect of trade with contracting parties not parties to such union or agreement shall not on the whole be higher or more restrictive than the general incidence of the duties and regulations of commerce applicable in the constituent territories prior to the formation of such union.”

## 4.1 Article XXIV restriction on external tariffs

We now isolate the implications of the pooling of market power under a CU from its relative inflexibility compared to an FTA. We do this by holding the external tariff of the CU constant at the pre-existing level  $t^*$ . This is well motivated experiment since the key WTO clause that sanctions PTAs, i.e. Article XXIV of GATT, forbids member countries of a PTA from raising tariffs on non-members. In our model, in the absence of such a restriction, CU members indeed raise their tariffs on outsiders due to the pooling of their market power whereas FTA members do not. What is the equilibrium outcome of the CU game when CU members are unable to raise their tariffs above the levels that prevail in the absence of any trade agreements?<sup>19</sup>

If CU members are prohibited from raising tariffs on non-members, the external tariffs under a CU and FTA are equal:

$$t_i(ij) = t_i(ij_u) = t^* \quad (34)$$

Given that they impose equal external tariffs and face the same tariffs in export markets, the welfare levels of a CU member is the same as that as an FTA member:  $w_m(ij_u) = w_m(ij)$  for  $m = i, j$ . It turns out that countries  $i$  and  $j$  still have an incentive to jointly deviate from their respective announcements  $\{j, k\}$  and  $\{i, k\}$  to the announcements  $\{j, \phi\}$  and  $\{i, \phi\}$  (converting free trade  $\langle F \rangle$  to a bilateral CU  $\langle ij_u \rangle$ ) even when the external tariffs are restricted by Article XXIV:

$$\Delta w_m(F - ij_u) < 0 \text{ for } m = i, j \quad (35)$$

As before, this deviation is self-enforcing and thus free trade fails to be coalition-proof under the restricted CU game. Thus, even when CU members are prevented from raising their tariffs on outsiders, a bilateral CU still emerges in coalition-proof Nash equilibria of the CU game. The only consequence of the tariff restriction is that it reduces the external tariff of the CU. Thus, though the restriction on external tariffs of CU members fails to further the cause of global free trade, it makes the resulting CU more attractive from a global welfare perspective by softening the adverse impact of the CU on the non-member

---

<sup>19</sup>Answering this question also allows us to highlight the value of flexibility enjoyed by FTA members since a CU whose external tariff equals that of an FTA differs from it in only one key respect: while FTA members can form independent FTAs with the non-member, CU members cannot. Thus, imposing the external tariff restriction of GATT Article XXIV allows us to utilize some relevant institutional features of the world trading system to isolate the value of the flexibility enjoyed by FTA members.

country.<sup>20</sup> Hereafter, we assume that a CU must abide by the external tariff restriction imposed by GATT Article XXIV. A useful implication of abiding by GATT Article XXIV in our model is that the welfare of members of a CU becomes equal to that of FTA members (since their external tariffs become equal and internal tariffs are zero in both cases). As a result, the two types of PTAs differ from one another only in terms of the flexibility available to members

## 4.2 The choice between FTAs and CUs

Thus far, we have required countries to form either FTAs or CUs without allowing them to pick between the two. What if both types of PTAs were permitted simultaneously in the same game? To address this question, consider the following PTA game. In the first stage, each country simultaneously announces the names of countries with whom it wants to sign a PTA as well as the nature of the PTA it seeks. For an agreement to arise between two countries, they need to not only announce each other's names but also call in favor of the same type of PTA (i.e. an FTA or a CU).

Country  $i$ 's strategy set  $\Omega_i$  consists of seven possible announcements:

$$\Omega_i = \{ \{ \phi, \phi \}, \{ j, \phi \}, \{ j_u, \phi \}, \{ \phi, k \}, \{ \phi, k_u \}, \{ j, k \}, \{ j_u, k_u \} \}$$

where  $\{ \phi, \phi \}$  denotes an announcement in favor of no FTA or CU with either trading partners,  $\{ j, \phi \}$  in favor of an FTA with only country  $j$ ;  $\{ j_u, \phi \}$  in favor of a CU with country  $j$ ;  $\{ \phi, k \}$  in favor of an FTA with only country  $k$ ;  $\{ \phi, k_u \}$  in favor of a CU with country  $k$ ;  $\{ j, k \}$  in favor of FTAs with both of them and  $\{ j_u, k_u \}$  in favor of a CU that includes all three countries which is tantamount to announcing in favor of global free trade. Here, it is important to re-emphasize that, due to the common external tariff requirement of a CU, a country cannot announce in favor of (i) two independent CUs with its trading partners or (ii) in favor of a bilateral CU and a bilateral FTA simultaneously. Finally, note that announcing  $\{ j_u, k_u \}$  is not equivalent to announcing in favor of a bilateral CU with either partner.

The mapping between announcements and trade agreements in the extended PTA game

---

<sup>20</sup>In a recent paper, Mrázová, et al. (2013) compare CU formation with and without the constraint on external tariffs imposed by Article XXIV. In their multi-country oligopoly model, holding constant the structure of the CU, the constraint imposed by Article XXIV improves world welfare. But this positive effect is counterbalanced by a negative compositional effect since Article XXIV encourages the formation of more symmetric CUs and thereby causes a greater volume of trade to be subject to tariffs.

follows naturally from our previous two games. To avoid redundancy, we simply note two novel aspects of the mapping between announcements and agreements in the extended game. First, no agreement  $\langle \Phi \rangle$  obtains in the PTA game whenever no announcements match or the only matching announcements are  $\{\phi, \phi\}$  where two announcements can fail to match just because two countries desire a different type of a PTA with each other. For example, no agreement  $\langle \Phi \rangle$  obtains when  $\sigma_i = \{j, \phi\}$ ;  $\sigma_j = \{i_u, \phi\}$ ; and  $\sigma_k = \{\varphi, \phi\}$ . In this case, although countries  $i$  and  $j$  announce each other's names, country  $i$  calls in favor of an FTA whereas country  $j$  announces in favor of a CU. No agreement obtains because countries  $i$  and  $j$  call in favor of different types of PTAs with each other and neither of them calls country  $k$ . The second point to note about the PTA game is that both of the following announcements map into global free trade:  $\{\sigma_i = \{j, k\}, \sigma_j = \{i, k\}, \sigma_k = \{i, j\}\}$  and  $\{\sigma_i = \{j_u, k_u\}, \sigma_j = \{i_u, k_u\}, \sigma_k = \{i_u, j_u\}\}$ . In other words, if all countries announce in favor of the same type of PTA with both their trading partners, the resulting outcome is free trade regardless of whether the PTA is an FTA or a CU. Other than these two aspects, the mappings between various announcements and trade agreements described in the FTA game and the CU game carry over naturally to the PTA game.

Combining the proofs of Propositions 2 and 4, we first argue that free trade does not arise in the CPNE of the PTA game. As before, it is immediate that countries  $i$  and  $j$  have an incentive to coalitionally deviate from their free trade announcements to the announcements  $\{j_u, \phi\}$  and  $\{i_u, \phi\}$  (thereby converting  $\langle F \rangle$  to  $\langle ij_u \rangle$ ) and it is a self-enforcing deviation since, taking country  $k$ 's announcement as fixed at  $\{i_u, j_u\}$ , neither of them has an incentive to further deviate unilaterally.

It is straightforward from part (i) of Lemmas 1 and 2 that  $\langle \Phi \rangle$  is not a CPNE since a pair of countries have an incentive to coalitionally deviate in order to form an FTA or a CU and such a joint deviation is self-enforcing. It is immediate from the inequalities reported in parts (iii) and (iv) of Lemma 1 that a member country and the non-member country facing a bilateral FTA have an incentive to engage in a coalitional deviation that would convert the FTA to a hub and spoke regime. Furthermore, this initial coalitional deviation is self-enforcing since no proper subset of the initially deviating countries has an incentive to further alter its announcement unilaterally. Therefore, announcement profiles that yield a bilateral FTA are not a CPNE of the PTA game.

Next, we show that a hub and spoke regime such as  $\langle ih \rangle$  also does not arise in the CPNE of the PTA game. Note from the inequality in part (v) of Lemma 1 that two spoke countries have incentives to coalitionally change their announcements, converting a hub

and spoke regime to free trade. This initial coalitional deviation is self-enforcing since the announcement profile supporting free trade is a Nash equilibrium. Therefore,  $\langle ih \rangle$  cannot arise in a CPNE. Finally, since member countries under a bilateral CU obtain the highest possible individual welfare, they have no incentive to unilaterally or coalitionally deviate, and thus, only those announcement profiles that yield  $\langle iju \rangle$  constitute a CPNE.

Based on our previous discussions of the FTA game and CU game, it is clear then that when countries are free to choose between the two types of PTAs, the CPNE of the PTA game yield  $\langle iju \rangle$ . One way of interpreting this result is that if the CU game is expanded to allow additional strategies that allow players to call in favor of an FTA, the result stated in Proposition 3B remains unaffected. This finding further supports the intuitive idea underlying the paper – i.e. owing to its relatively inflexible nature, a CU is a more coalitionally stable agreement compared to an FTA. However, it also seems fair to acknowledge that, while our model does a nice job of capturing this fundamental difference between FTAs and CUs, we do not think that the PTA game analyzed above captures some important real-world factors that are likely to determine the choice between FTAs and CUs. For example, by design, the model abstracts from the difficulties that actual CU members face in reaching agreement on and then enforcing a common external tariff. Similarly, the sharing of tariff revenue among CU members does not occur as seamlessly in the real world as is assumed in the model. Finally, the underlying trading structure assumed by us is one in which the exclusion incentive of a CU ends up driving equilibrium outcomes. While such an incentive is likely to arise in other trading structures, its magnitude may not be large enough to yield a CU as an equilibrium outcome. For example, in the competing exporters framework of Saggi et al. (2013), the exclusion incentive of a CU is weaker and the equilibrium outcome of the CU game is global free trade.

### 4.3 Implications of asymmetry: an illustration

Thus far, we have assumed that countries are symmetric in every respect except for the fact that each country has a comparative advantage in a unique good. It is natural to wonder whether our key results extend to a setting where countries differ from one another in some respect. To gain some insight into how asymmetry might matter, it proves instructive to focus on the following case: suppose two countries (denoted by  $l$  and  $l'$ ) have larger market demands than the third (denoted by  $s$ ). More specifically, let the pattern of demand asymmetry be captured by the demand intercept  $\alpha$  (see equation 2) in the following fashion:  $\alpha_s = 1 < \alpha_l = \alpha_{l'} = \alpha$ . In order to guarantee interior solutions to the optimal tariff

problems of all countries, we assume that  $\alpha < \bar{\alpha}(\lambda) = 1 + \lambda/4$ .

How does a country's incentive to form a bilateral PTA with another depend on the distribution of market sizes across countries? We address this key question in the following lemma:

**Lemma 3:** *Let country  $j$  be a PTA partner of country  $i$  under regime  $r$  but not regime  $v$ , where  $r, v = \langle \Phi \rangle, \langle ij \rangle, \langle ij_u \rangle, \langle ih \rangle$  or  $\langle F \rangle$  and  $r \neq v$ . Then, the following holds:  $\frac{\partial \Delta w_i(r-v)}{\partial \alpha_i} < 0 < \frac{\partial \Delta w_i(r-v)}{\partial \alpha_j}$ .*

First, observe that the countries with larger markets have greater market power, and therefore gain more from their optimal tariffs than the smaller country. The intuition underlying the  $\frac{\partial \Delta w_i(r-v)}{\partial \alpha_i} < 0$  is as follows. Due to the smaller volume of their exports, countries with larger markets benefit less from tariff reductions granted by the smaller country. Similarly, larger countries have relatively more to lose from eliminating their own optimal tariffs since these tariffs apply to relatively larger import volumes (or to relatively inelastic export supply curves). Thus, a country's willingness to enter into a bilateral trade agreement with another depends negatively on its own market size.

A similar intuition underlies the other inequality (i.e.  $\frac{\partial \Delta w_i(r-v)}{\partial \alpha_j} > 0$ ). The larger the market size of a country's partner, the larger the increase in its export surplus from the elimination of its partner's optimal tariff and the smaller the loss due to its own trade liberalization since the tariff reduction applies to a smaller volume of imports (due to the larger market size of its partner). The two inequalities reported in Lemma 3 imply that a country prefers to form a bilateral PTA with the larger of its two trading partners:

$$w_i(ij_u) = w_i(ij) \geq w_i(ik) = w_i(ik_u) \text{ iff } \alpha_j \geq \alpha_k. \quad (36)$$

Let  $\alpha_i(r-v)$  denote the critical threshold of market size asymmetry at which country  $i$  is indifferent between regimes  $r$  and  $v$  so that  $\Delta w_i(r-v) = 0$ . For example, consider the following two regimes:  $\langle ll' \rangle$  and  $\langle lh \rangle$  from the perspective of countries  $s$  and  $l$ . Using the above lemma, we know that (i)  $\frac{\partial \Delta w_s(lh-l')}{\partial \alpha_s} < 0$  and (ii)  $\frac{\partial \Delta w_l(lh-l')}{\partial \alpha_s} > 0$ . We know from the discussion under symmetry that a country prefers being a spoke to being a non-member facing an FTA ( $\Delta w_s(lh-l') > 0$ ). As country  $s$  becomes smaller in size relative to the other two countries, using (i) we can show that  $w_s(lh-l') > 0$  for all  $\alpha$  – i.e. the result that it is better to be a spoke than to be an outsider extends to the asymmetric setting for the smaller country. Furthermore, there exists a critical threshold of market size asymmetry ( $\alpha_l(lh-l')$ ) at which country  $l$  is indifferent between regimes  $\langle ll' \rangle$  and  $\langle lh \rangle$ . Recall that under symmetry a country prefers being a hub to being a member under an FTA. Using

(ii), we can show that the incentive to become a hub decreases as country  $s$  becomes smaller in size relative to the other two countries. At the limit when  $\alpha = \bar{\alpha}(\lambda)$ , we have  $\Delta w_l(lh - ll') < 0$  and thus it follows that country  $l$  prefers  $\langle ll' \rangle$  to  $\langle lh \rangle$  when  $\alpha > \alpha_l(lh - ll')$ . Using this logic and utilizing the results under symmetry as well as Lemma 3, we prove the following proposition, illustrated in figure 1, in the appendix:

**Proposition 4:** *Let  $\alpha_s = 1 < \alpha_l = \alpha_{l'} = \alpha < \bar{\alpha}(\lambda)$ . Then, the following hold:*

(i) **FTA game:** *When  $\alpha < \alpha(mh - ll')$  for  $m = l, l'$ , the strategy profile that yields global free trade  $\langle F \rangle$  is the unique CPNE of the FTA game whereas when  $\alpha > \alpha(mh - ll')$  those announcement profiles that yield the bilateral FTA between the two larger countries  $\langle ll' \rangle$  constitute the set of CPNE.<sup>21</sup>*

(ii) **CU game:** *Only those announcement profiles that yield the bilateral CU between the two larger countries  $\langle ll'_u \rangle$  constitute the CPNE of the CU game.<sup>22</sup>*

Insert Figure 1

The intuition behind the first statement in part (i) of Proposition 4 is very similar to Proposition 2. The two larger countries benefit from excluding the smaller country from their trade agreement:  $\Delta w_l(F - ll') = \Delta w'_l(F - ll') < 0$ . Thus, taking the announcement of country  $s$  as fixed, countries  $l$  and  $l'$  have an incentive to jointly deviate from  $\sigma_l = \{l', s\}$  and  $\sigma_{l'} = \{l, s\}$  to  $\sigma_l = \{l', \phi\}$  and  $\sigma_{l'} = \{l, \phi\}$  respectively. However, this joint deviation turns out to be not self-enforcing when the degree of market size asymmetry is sufficiently small ( $\alpha < \alpha_l(lh - ll')$ ) since each larger country, taking the announcement of the smaller country fixed at  $\sigma_s = \{l, l'\}$ , has an incentive to deviate from its announcement under the coalitional deviation by altering its announcement to include the smaller country's name in order to position itself as a hub. For example, given  $\sigma_{l'} = \{l, \phi\}$  and  $\sigma'_k = \{l, l'\}$ , country  $l$  has an incentive to alter its announcement from  $\sigma_l = \{l', \phi\}$  to  $\sigma_l = \{l, s\}$  so as to obtain the hub and spoke arrangement  $\langle lh \rangle$ . Thus the result stated in Proposition 3A does not require symmetry; it continues to hold as long as the degree of demand asymmetry between countries is not too pronounced.

When the market size of two larger countries is sufficiently large relative to the smaller country (i.e.  $\alpha > \alpha_l(lh - ll')$ ), the degree of external trade diversion caused by the bilateral FTA  $\langle ll' \rangle$  is high and this tends to lower the optimal tariffs of the smaller country.

---

<sup>21</sup>All strategy profiles with the following two properties yield  $\langle ll' \rangle$ : (i)  $l \in \sigma_{l'}$  and  $l' \in \sigma_l$  and (ii)  $s \notin \sigma_l$  and  $s \notin \sigma_{l'}$ .

<sup>22</sup>All strategy profiles with the following two properties yield  $\langle ll'_u \rangle$ : (i)  $l^u \in \sigma_{l'}$  and  $l'_u \in \sigma_l$  and (ii)  $s^u \notin \sigma_l$  and/or  $s^u \notin \sigma_{l'}$ .

Therefore, from the perspective of each larger country, starting at the bilateral FTA  $\langle ll' \rangle$  the benefit of an additional FTA with the smaller country is limited while the domestic surplus loss incurred by giving up its own optimal tariff is sizeable. Under such a scenario, each larger country prefers to be the member of the bilateral FTA  $\langle ll' \rangle$  relative to becoming a hub with two independent FTAs. As a result, the joint announcement deviation of the larger countries from their announcements  $\sigma_l = \{l', s\}$  and  $\sigma_{l'} = \{l, s\}$  to  $\sigma_l = \{l', \phi\}$  and  $\sigma_{l'} = \{l, \phi\}$  is self-enforcing and free trade fails to be a CPNE. Thus, when the degree of demand asymmetry exceeds the critical threshold  $\alpha(lh - ll')$ , the announcement profile  $\{\sigma_l = \{l', \phi\}, \sigma_{l'} = \{l, \phi\}, \sigma_s = \{\phi, \phi\}\}$  leading to the bilateral FTA  $\langle ll' \rangle$  is the unique CPNE of the FTA game.

Finally, consider part (ii) of Proposition 4, which states that the result under symmetry (Proposition 3B) extends without any modifications to the asymmetric setting examined here. As in the FTA game, the two larger countries benefit from excluding the smaller country – i.e.  $\Delta w_l(F - ll'_u) = \Delta w_{l'}(F - ll'_u) < 0$  – and the inability of a CU member to enter into an independent trade agreement with the outsider implies that the exclusion incentive is exercised in equilibrium. Since the larger countries prefer the bilateral CU  $\langle ll'_u \rangle$  to no agreement  $\langle \Phi \rangle$ , there exist no coalitional deviation incentive on their part from the announcements that support  $\langle ll'_u \rangle$ .

The analysis above shows that the lure of a hub and spoke arrangement that undermines a bilateral FTA is not as strong when the excluded country is sufficiently small relative to the other two countries. This is a fairly intuitive result and it provides useful perspective on when and why the pursuit of bilateral FTAs does not necessarily end up in global free trade. It also fits well with related existing literature that generally finds free trade to be more likely to obtain when countries are relatively symmetric to one another.

## 5 Conclusion

At a time when multilateral trade liberalization at the WTO seems to have come to a grinding halt, the question whether PTAs have a useful role to play in the global trading system seems to have acquired a greater degree of urgency. After all, other than unilateral trade liberalization, over the last decade or so PTAs appear to be the only game in town for countries interested in undertaking reciprocal trade liberalization. Of course, the concern that the pursuit of PTAs might undermine whatever appetite remains in the global trading system for multilateral trade liberalization remains a real one.

In this paper, we attempt to isolate the implications of the two commonly occurring PTAs, i.e. FTAs and CUs, for the prospects of global free trade. The starting point of our analysis is a well understood difference between these two types of PTAs: while CU members impose common external tariffs, FTA members are free to implement external tariffs of own choosing. This difference in tariff setting behavior between the two PTAs implies that while FTA members have less market power than CU members, they enjoy a greater degree of flexibility in the sense that they are free to enter into further trade agreements with outsiders whereas CU members can only do so with consent of existing members.

An important result of this paper is that the formation of either type of PTA induces the non-member to lower its tariffs on PTA members: the external trade diversion – i.e. the reduction in the volume of exports flowing from PTA members to the non-members – reduces the ability of the non-member to manipulate its terms of trade thereby making it optimal for it to lower its tariffs. Thus, PTA formation not only benefits members because of the internal trade liberalization that they undertake but also because of the external trade liberalization induced abroad.

The central result of the paper is that the more flexible nature of FTAs (which in turn emanates from independent external tariff setting on the part of FTA members) helps in the attainment of global free trade. Specifically, in the CPNE of the FTA game, each country ends up entering into FTAs with both its trading partners, an outcome under which all countries impose zero tariffs on each other. By contrast, a bilateral CU is the outcome of all announcement profiles that constitute the CPNE of the CU game. We show that this result holds even when demand, and therefore market power, is asymmetric across countries. This result captures the intuition that, once formed, CUs constrain their members in a way that FTAs do not. Finally, we also find that the tariff restriction imposed by Article XXIV of the GATT fails to further the cause of global free trade although it makes the resulting CU more attractive from a global welfare perspective. Thus, our model shows that while the two key attributes of a CU – i.e. the pooling of market power by its members and the inability of any one member to enter into independent trade agreements with outsiders – both impose welfare costs on the world trading system, the latter is a more important obstacle in path of global free trade.

## 6 Appendix

### 6.1 Supporting calculations

Here we report the key formulae that are necessary for proving our results. For an arbitrary tariff vector  $\mathbf{t} = (t_{ij}, t_{ik}, t_{ji}, t_{jk}, t_{ki}, t_{kj})$ , we can write country  $i$ 's welfare as

$$w_i = CS_i + PS_i + TR_i$$

where consumer surplus in country  $i$  equals

$$CS_i = \frac{1}{2} [(\alpha_i - p_i^I)^2 + (\alpha_i - p_j^J - t_{ij})^2 + (\alpha_i - p_k^K - t_{ik})^2],$$

whereas its producer surplus equals

$$PS_i = \frac{(\lambda + 1)}{2} (p_i^I)^2 + \frac{1}{2} [(p_j^J + t_{ij})^2 + (p_k^K + t_{ik})^2]$$

and the tariff revenue is given by

$$TR_i = t_{ij}(2t_{kj} - \alpha_k) + t_{ik}(2t_{jk} - \alpha_j) + \sum_{z \neq i} t_{iz}(4 + \lambda_z)p_z^Z - \alpha_z$$

Equilibrium prices are given by

$$p_i^I = \frac{\sum_{z=i,j,k} \alpha_z - 2 \sum_{z \neq i} t_{zi}}{\lambda + 6}; p_j^J = \frac{\sum_{z=i,j,k} \alpha_z - 2 \sum_{z \neq j} t_{zj}}{\lambda + 6}; \text{ and } p_k^K = \frac{\sum_{z=i,j,k} \alpha_z - 2 \sum_{z \neq k} t_{zk}}{\lambda + 6}$$

Finally, optimal tariffs under each trade regime are given by:

$$t_{ll'}^* = t_{ll'}(sl) = \frac{\alpha\lambda^2 + 8\alpha\lambda + 8\alpha - 2\lambda - 8}{(\lambda + 6)(\lambda^2 + 12\lambda + 28)}$$

$$t_{ls}^* = t_{ls}(ll') = \frac{\alpha\lambda + 2\alpha - 2}{(\lambda^2 + 12\lambda + 28)}$$

$$t_{sl}^* = t_{sl}(sl') = \frac{\lambda^2 + 10\lambda + 20 - 4\alpha(\lambda + 5)}{(\lambda + 6)(\lambda^2 + 12\lambda + 28)}$$

$$t_{sl}(ll') = t_{sl}(l'h) = \frac{\lambda + 4(1 - \alpha)}{(\lambda + 4)(\lambda + 8)}$$

$$t_{ll'}(sl') = t_{ll'}(sh) = \frac{\alpha(\lambda + 2) - 2}{(\lambda + 4)(\lambda + 8)}$$

Note that  $t_{sl}(ll') = t_{sl}(lh) \geq 0$  only when  $\alpha < \bar{\alpha}(\lambda) = 1 + \lambda/4$ . Using the above formulae and the optimal tariff levels reported in the text, we can easily calculate welfare levels under all possible trade agreements for both cases of symmetry and asymmetry. To save space, we do not include the algebraic details underlying these straightforward calculations. The proofs of Lemmas 1-3 follow immediately from these welfare comparisons.

## 6.2 Proof of Proposition 2A

In the text, we showed that parsimonious announcement profiles leading to  $\langle \Phi \rangle$ ,  $\langle ij \rangle$ , and  $\langle ih \rangle$  are Nash equilibria of the FTA game. We next consider non-parsimonious announcement profiles and show that these are not Nash equilibria.

First note that, in any non-parsimonious announcement profile leading to no agreement  $\langle \Phi \rangle$ , it must be the case that countries' announcements do not match with each other. Under such a case, since countries always benefit from FTA formation (from Lemma 1), a country whose name is announced has an incentive to unilaterally change its announcement, leading to a bilateral FTA. Therefore, any non-parsimonious announcement profile that leads to  $\langle \Phi \rangle$  cannot be a Nash equilibrium.

Now consider the non-parsimonious announcement profiles leading to a bilateral FTA  $\langle ij \rangle$ . First, consider the following scenario:  $\sigma_i = \{j, \phi\}$  and  $\sigma_j = \{i, \phi\}$  while  $\sigma_k = \{i, \phi\}$  or  $\{\phi, j\}$  or  $\{i, j\}$ . Under such a case, depending on country  $k$ 's announcement, either country  $i$  or country  $j$  has an incentive to unilaterally change its announcement to include  $k$  in order to become a hub under a hub and spoke regime. Second, suppose (a)  $\sigma_i = \{j, k\}$  and  $\sigma_j = \{i, \phi\}$  but  $i \notin \sigma_k$  or (b)  $\sigma_i = \{j, \phi\}$  and  $\sigma_j = \{i, k\}$  but  $j \notin \sigma_k$  or (c)  $\sigma_i = \{j, k\}$  and  $\sigma_j = \{i, k\}$  but  $i, j \notin \sigma_k$ . Under all three cases, since country  $k$  prefers being spoke to being a non-member under a bilateral FTA (from Lemma 1), it has an incentive to unilaterally change its announcement in order to become a spoke under a hub and spoke regime. As a result, any non-parsimonious announcement leading to a bilateral FTA  $\langle ij \rangle$  cannot be a Nash equilibrium.

Finally, suppose that  $\sigma_i = \{j, k\}, \sigma_j = \{i, k\}, \sigma_k = \{i, \phi\}$  or  $\sigma_i = \{j, k\}, \sigma_j = \{i, \phi\}, \sigma_k = \{i, j\}$ . Under both scenarios, the spoke country that does not announce the other spoke has an incentive to unilaterally change its announcement to also call the other spoke which leads to free trade  $\langle F \rangle$ . Therefore, the initial announcement profiles are not Nash equilibrium.

### 6.3 Proof of Proposition 2B

In the text, we showed that parsimonious announcement profile leading to  $\langle \Phi \rangle$  is a Nash equilibrium of the CU game. We next consider non-parsimonious announcement profiles and show that these are not Nash equilibria. As in the FTA game, in any non-parsimonious announcement profile leading to no agreement  $\langle \Phi \rangle$ , it must be the case that countries' announcements do not match with each other. Under such a case, since countries always benefit from CU formation (from Lemma 2), a country whose name is announced has an incentive to unilaterally change its announcement, leading to a bilateral CU. Therefore, any non-parsimonious announcement profile that leads to  $\langle \Phi \rangle$  cannot be a Nash equilibrium.

Next, we show that the following announcement profile that yields a bilateral CU  $\langle ij_u \rangle$ :  $\sigma_i = \{j_u, k_u\}, \sigma_j = \{i_u, k_u\}, \sigma_k = \{\phi, \phi\}$  is not a Nash equilibrium. Under such a case, since country  $k$  prefers free trade to being a non-member under  $\langle ij_u \rangle$  (from Lemma 2), it has an incentive to unilaterally change its announcement from  $\{\phi, \phi\}$  to  $\{i_u, j_u\}$  that would convert  $\langle ij_u \rangle$  to  $\langle F \rangle$ .

### 6.4 Proof of Proposition 3A

We begin by considering whether the announcement profile  $\Omega^F$  that leads to  $\langle F \rangle$  is a CPNE. Since world welfare is the highest under  $\langle F \rangle$  and countries have equal welfare due to symmetry, each country is better off under  $\langle F \rangle$  relative to  $\langle \Phi \rangle$  and thus we can immediately rule out any coalitional announcement deviations that would lead to  $\langle \Phi \rangle$  replacing  $\langle F \rangle$ . Similarly, we know that no two countries (say  $j$  and  $k$ ) have incentives to jointly alter their announcements from  $\{i, k\}$  to  $\{i, \phi\}$  and  $\{i, j\}$  to  $\{i, \phi\}$ , respectively since doing so would lead to a deviation from  $\langle F \rangle$  to  $\langle ih \rangle$  where both become spokes (and spokes are worse off relative to free trade). Finally, taking the announcement of their complement (country  $k$ ) as fixed, consider the joint deviation of two countries (say  $i$  and  $j$ ) from their announcements  $\{j, k\}$  and  $\{i, k\}$  to  $\{j, \phi\}$  and  $\{i, \phi\}$  respectively. This joint deviation implies a coalitional deviation from  $\langle F \rangle$  to  $\langle ij \rangle$ . From Lemma 1 part (ii) we know that  $\Delta w_m(F - ij) < 0$ . This implies that, taking the announcement of their complement (country  $k$ ) fixed at  $\sigma_k = \{i, j\}$ , the above coalitional deviation in announcements would occur. The question then becomes whether this joint deviation is self-enforcing. We next argue that *it is not*. This is because given that  $\sigma_k = \{i, j\}$ , we know from Lemma 1 part (iii) that country  $i$  has an incentive to alter its announcement  $\{j, \phi\}$  to  $\{j, k\}$  in order to create the trading regime  $\langle ih \rangle$  where it becomes the hub. Similarly, country  $j$  has an incentive to alter its announcement  $\{i, \phi\}$  to  $\{i, k\}$  so as to itself become the hub. Thus, the initial coalitional deviation that can cause  $\langle F \rangle$  to be replaced by  $\langle ij \rangle$  is not self-enforcing. As a result, the announcement profile  $\Omega^F$  is a CPNE.

Next, we consider whether the announcement profile  $\Omega^\Phi$  is a CPNE. Note from Lemma 1

part (i) that starting at  $\Omega^\Phi$  any two countries (say  $i$  and  $j$ ) have an incentive to coalitionally change their announcements from  $\{\phi, \phi\}$  and  $\{\phi, \phi\}$  to  $\{j, \phi\}$  and  $\{i, \phi\}$  respectively, taking country  $k$ 's announcement fixed:  $\sigma_k = \{\phi, \phi\}$ . This initial deviation is self-enforcing since no proper subset of the initially deviating countries (neither  $i$  nor  $j$ ) has an incentive to alter its announcement unilaterally (i.e. announcement profile that leads to  $\langle ij \rangle$  is a Nash equilibrium). Therefore, the announcement profile that leads to  $\langle \Phi \rangle$  is not a CPNE.

We next show that  $\Omega^{ij}$  is not a CPNE. From parts (iii) and (iv) of Lemma 1 we know that starting at  $\Omega^{ij}$  countries  $i$  and  $k$  have an incentive to coalitionally change their announcements from  $\{j, \phi\}$  and  $\{\phi, \phi\}$  to  $\{j, k\}$  and  $\{i, \phi\}$  respectively, taking country  $j$ 's announcement fixed at  $\sigma_j = \{i, \phi\}$ . This initial coalitional deviation would convert  $\langle ij \rangle$  to  $\langle ih \rangle$  where  $i$  is the hub and  $j$  and  $k$  are spokes. Furthermore, this initial coalitional deviation is self-enforcing since no proper subset of the initially deviating countries (neither  $i$  nor  $k$ ) has an incentive to unilaterally alter its announcement since the announcement profile  $\Omega^{ih}$  that leads to  $\langle ih \rangle$  is a Nash equilibrium. Therefore, the announcement profile  $\Omega^{ij}$  is not a CPNE.

Finally, we show that  $\Omega^{ih}$  is not a CPNE. Note from Lemma 1 part (v) that starting at  $\Omega^{ih}$  countries  $j$  and  $k$  have an incentive to coalitionally change their announcements from  $\{i, \phi\}$  and  $\{i, \phi\}$  to  $\{i, k\}$  and  $\{i, j\}$  respectively, taking country  $i$ 's announcement fixed at  $\sigma_i = \{j, k\}$ . This coalitional deviation converts the trade regime from  $\langle ih \rangle$  to  $\langle F \rangle$ . Furthermore, this initial coalitional deviation is self-enforcing since no proper subset of the initially deviating countries (neither  $j$  nor  $k$ ) has an incentive to deviate unilaterally since  $\Omega^F$  is a Nash equilibrium. Therefore, the announcement profile  $\Omega^{ih}$  is not a CPNE.

## 6.5 Proof of Proposition 4

Note first that, since the external tariff restriction of Article XXIV is assumed to be binding, the welfare level of a member of a bilateral FTA is the same as that of member of a CU, i.e.  $w_i(ij) = w_i(ij^u)$  for all  $i, j = l, l', s$ . Using the welfare levels reported above, we can

easily show the following inequalities:

- (i)  $\Delta w_l(ll' - \Phi) > 0$  for all  $\alpha$
- (ii)  $\Delta w_l(ll' - sl) = \Delta w_l(ll'_u - sl'_u) > 0$  for all  $\alpha$
- (iii)  $\Delta w_s(lh - ll') > 0$  for all  $\alpha$  while  $\Delta w_l(lh - ll') > 0$  when  $\alpha < \alpha_l(lh - ll')$
- (iv)  $\Delta w_l(sh - ll') < 0 < \Delta w_s(sh - ll')$  for all  $\alpha$
- (v)  $\Delta w_l(F - ll') = \Delta w_l(F - ll'_u) < 0 < \Delta w_s(F - ll') = \Delta w_s(F - ll'_u)$  for all  $\alpha$
- (vi)  $\Delta w_s(sl - \Phi) > 0$  for all  $\alpha$  while  $\Delta w_l(sl - \Phi) > 0$  when  $\alpha < \alpha_l(sl - \Phi)$
- (vii)  $\Delta w_l(lh - sl) > 0$  and  $\Delta w_{l'}(lh - sl) > 0$  for all  $\alpha$
- (viii)  $\Delta w_s(sh - sl') > 0$  for all  $\alpha$  while  $\Delta w_l(sh - sl') > 0$  when  $\alpha < \alpha_l(sh - sl')$
- (ix)  $\Delta w_s(F - \Phi) > 0$  for all  $\alpha$  while  $\Delta w_l(F - \Phi) > 0$  when  $\alpha < \alpha_l(F - \Phi)$
- (x)  $\Delta w_s(F - sl) > 0$  for all  $\alpha$  while  $\Delta w_l(F - \Phi) > 0$  when  $\alpha < \alpha_l(F - sl)$
- (xi)  $\Delta w_l(F - sh) > 0$  for all  $\alpha$
- (xii)  $\Delta w_s(F - l'h) > 0$  for all  $\alpha$  while  $\Delta w_l(F - l'h) > 0$  when  $\alpha < \alpha_l(F - l'h)$
- (xiii)  $\Delta w_s(F - ll') > 0$  and  $\Delta w_l(F - sl') > 0$  for all  $\alpha$

We first consider the FTA game and prove Proposition 4(i).

#### CPNE of the FTA game under asymmetry

To avoid redundancy, we focus on parsimonious Nash announcements. Our proof extends to the case of non-parsimonious Nash announcements (which can arise in the FTA game under asymmetry).

First note from inequality (i) that starting at  $\Omega^\Phi = \{\sigma_s = \{\phi, \phi\}, \sigma_l = \{\phi, \phi\}, \sigma_{l'} = \{\phi, \phi\}\}$  the two large countries  $l$  and  $l'$  have an incentive to coalitionally change their announcements from  $\{\phi, \phi\}$  to  $\{l', \phi\}$  and  $\{l, \phi\}$  respectively, taking country  $s$ 's announcement fixed at  $\{\phi, \phi\}$ . This initial deviation is self-enforcing since no proper subset of the initially deviating countries (neither  $l$  nor  $l'$ ) has an incentive to alter its announcement unilaterally. Therefore, the announcement profile  $\Omega^\Phi$  that leads to  $\langle \Phi \rangle$  is not a CPNE.

We next show that the announcement profile  $\Omega^{sl} = \{\sigma_s = \{l, \phi\}, \sigma_l = \{\phi, s\}, \sigma_{l'} = \{\phi, \phi\}\}$  that leads to  $\langle sl \rangle$  is not a CPNE. Note first that  $\alpha_l(sh - sl') > \alpha_l(sl - \Phi)$  for all  $\lambda$ . From part (vi) we know that, when  $\alpha > \alpha_l(sl - \Phi)$  country  $l$  has an incentive to unilaterally change its announcement from  $\{\phi, s\}$  to  $\{\phi, \phi\}$  and thus the announcement profile  $\Omega^{sl}$  is not even a Nash equilibrium when  $\alpha > \alpha_l(sl - \Phi)$ . Moreover, we know from part (viii) that, when  $\alpha < \alpha_l(sh - sl')$ , countries  $s$  and  $l$  have an incentive to coalitionally change their announcements from  $\{l, \phi\}$  and  $\{\phi, \phi\}$  to  $\{l, l'\}$  and  $\{\phi, s\}$  respectively, taking country  $l'$ 's announcement fixed at  $\{\phi, s\}$ . This initial coalitional deviation would convert  $\langle sl' \rangle$  to  $\langle sh \rangle$  and it is self-enforcing in nature. Since  $\alpha_l(sh - sl') > \alpha_l(sl - \Phi)$  for all  $\lambda$ , it is immediate

that the announcement profile  $\Omega^{sl}$  is not a CPNE.

Consider now the announcement profile  $\Omega^{sh} = \{\sigma_s = \{l, l'\}, \sigma_l = \{\phi, s\}, \sigma_{l'} = \{\phi, s\}\}$ . We know from part (xi) that countries  $l$  and  $l'$  have an incentive to coalitionally change their announcements from  $\{\phi, s\}$  and  $\{\phi, s\}$  to  $\{l', s\}$  and  $\{l, s\}$  respectively, taking country  $s$ 's announcement fixed at  $\{l, l'\}$ . This initial coalitional deviation would convert  $\langle sh \rangle$  to  $\langle F \rangle$ . From parts (xii) and (xiii), we know that, when  $\alpha < \alpha_l(F - l'h)$ , this coalitional deviation is self-enforcing since neither large country has an incentive to unilaterally alter its announcement. Moreover, we know from part (iv) that countries  $l$  and  $l'$  have an incentive to coalitionally change their announcements from  $\{\phi, s\}$  and  $\{\phi, s\}$  to  $\{l', \phi\}$  and  $\{l, \phi\}$  respectively, taking country  $s$ 's announcement fixed at  $\{l, l'\}$ . This initial coalitional deviation would convert  $\langle sh \rangle$  to  $\langle ll' \rangle$ . From part (iii) we know that this coalitional deviation is self-enforcing when  $\alpha > \alpha_l(lh - ll')$ . Since  $\alpha_l(lh - ll') < \alpha < \alpha_l(F - l'h)$  for all for all  $\lambda$ , the announcement profile that  $\Omega^{sh}$  is not a CPNE.

We next consider the announcement profile  $\Omega^{lh} = \{\sigma_s = \{\phi, l'\}, \sigma_l = \{l', \phi\}, \sigma_{l'} = \{l, s\}\}$ . Note from part (iii) that  $\Delta w_l(lh - ll') < 0$  when  $\alpha > \alpha_l(lh - ll')$ . Thus, under such a case, country  $l$  has an incentive to unilaterally change its announcement from  $\{l', \phi\}$  to  $\{\phi, \phi\}$  and thus  $\Omega^{lh}$  is not even a Nash equilibrium when  $\alpha > \alpha_l(lh - ll')$ . Moreover, we know from part (xii) that, when  $\alpha < \alpha_l(F - l'h)$ , countries  $s$  and  $l$  have an incentive to coalitionally change their announcements from  $\{\phi, l'\}$  and  $\{l', \phi\}$  to  $\{l', s\}$  and  $\{l, l'\}$  respectively, taking country  $l'$ 's announcement fixed at  $\{l, s\}$ . This initial coalitional deviation would convert  $\langle lh \rangle$  to  $\langle F \rangle$  and it self-enforcing in nature. Given that we have  $\alpha_l(F - l'h) > \alpha_l(lh - ll')$  for all  $\lambda$ , it follows that the announcement profile  $\Omega^{lh}$  is not a CPNE.

It is immediate from the above discussion that the only remaining candidates for being a CPNE under the FTA game are the announcement profiles  $\Omega^{ll'}$  and  $\Omega^F$ . We first consider the announcement profile  $\Omega^{ll'} = \{\sigma_s = \{\phi, \phi\}, \sigma_l = \{l', \phi\}, \sigma_{l'} = \{l, \phi\}\}$ . On one hand, we know from parts (i), (ii), (iv) and (v) that there exist no unilateral or coalitional announcement deviations that alters the trade regime from  $\langle ll' \rangle$  to  $\langle \Phi \rangle$ ,  $\langle sl \rangle$ ,  $\langle sh \rangle$  and  $\langle F \rangle$ . On the other hand, it is immediate from part (iii) that, when  $\alpha < \alpha_l(lh - ll')$ , countries  $l$  and  $s$  have an incentive to change their respective announcement profiles  $\{l', \phi\}$  and  $\{\phi, \phi\}$  to  $\{l', s\}$  and  $\{l, \phi\}$  respectively, taking country  $l'$ 's announcement fixed at  $\sigma_{l'} = \{l, \phi\}$ . This initial coalitional deviation would convert  $\langle ll' \rangle$  to  $\langle lh \rangle$  and it is self-enforcing in nature. As a result, the announcement profile that leads to  $\langle ll' \rangle$  is a CPNE when  $\alpha \geq \alpha_l(lh - ll')$ .

We now show when and why  $\Omega^F$  is a CPNE of the FTA game. Note from part (v) that starting at  $\Omega^F = \{\sigma_s = \{l, l'\}, \sigma_l = \{l', s\}, \sigma_{l'} = \{l, s\}\}$  countries  $l$  and  $l'$  have an incentive to coalitionally change their announcements from  $\{l', s\}$  and  $\{l, s\}$  to  $\{l', \phi\}$  and  $\{l, \phi\}$  respectively, taking country  $s$ 's announcement fixed at  $\{l, l'\}$ . This initial coalitional deviation would convert  $\langle F \rangle$  to  $\langle ll' \rangle$ . From part (iii), we know that the initial coalitional deviation is self-enforcing only when  $\alpha > \alpha_l(lh - ll')$ . Note also that when  $\alpha \leq \alpha_l(lh - ll')$  there exist no unilateral or self-enforcing coalitional announcement deviations from  $\Omega^F$

since  $\alpha_l(lh - ll') < \alpha_l(F - \Phi)$  and  $\alpha_l(lh - ll') < \alpha_l(F - l'h)$ . As a result, the announcement profile  $\Omega^F$  is a CPNE when  $\alpha \leq \alpha_l(lh - ll')$ .

We now prove Proposition 4(ii).

### CPNE of the CU game under asymmetry

It is immediate from parts (i), (ii) and (v) that there exist no unilateral or coalitional announcement deviations from the various announcement profiles that lead to  $\langle ll'_u \rangle$ . Therefore, all announcement profiles leading to  $\langle ll'_u \rangle$  are not just CPNE of the CU game but also constitute its strong Nash equilibria. Moreover the announcement profiles that lead to  $\langle \Phi \rangle$ ,  $\langle sl_u \rangle$ , and  $\langle F \rangle$  fail to be CPNE since the coalitional announcement deviations to the announcement profiles that lead to  $\langle ll'_u \rangle$  occur and they are self-enforcing in nature.

## References

- [1] Bagwell, Kyle and Robert. W. Staiger, 1997a. "Multilateral Tariff Cooperation During the Formation of Customs Unions." *Journal of International Economics*, Elsevier 42, 91-123.
- [2] Bagwell, Kyle and Robert. W. Staiger, 1997b. "Multilateral Tariff Cooperation During the Formation of Free Trade Areas." *International Economic Review* 38, 291-319.
- [3] Bagwell, Kyle and Robert. W. Staiger, 1998. "Regionalism and Multilateral Tariff Cooperation." In John Piggott and Alan Woodland, eds, *International Trade Policy and the Pacific Rim*, London: MacMillan.
- [4] Bernheim, Douglas B., Bezalel Peleg and Michael Whinston, 1987. "Coalition-proof Nash Equilibria I. Concepts." *Journal of Economic Theory* 42, 1-12.
- [5] Bond, Eric W., and Constantinos Syropoulos, 1996. "The Size of Trading Blocs: Market Power and World Welfare Effects." *Journal of International Economics* 40, 411-437.
- [6] Bond, Eric W., Raymond G. Riezman, and Constantinos Syropoulos, 2004. "A Strategic and Welfare Theoretic Analysis of Free Trade Areas." *Journal of International Economics* 64, 1-27.
- [7] Broda, Christian, Nuno Limao and David E. Weinstein, 2008. "Optimal Tariffs and Market Power: The Evidence," *American Economic Review*, 98(5), 2032-65.
- [8] Coates, Daniel and Rodney Ludema, 2001. "A Theory of Trade Policy Leadership." *Journal of Development Economics* 65, 1-29.

- [9] Estevadeordal, Antoni, Caroline Freund and Emanuel Ornelas, 2008. “Does Regionalism Affect Trade Liberalization toward Non Members?” *Quarterly Journal of Economics* 123(4), 1531-1575.
- [10] Feenstra, Robert, 2004. *Advanced International Trade*. Princeton University Press.
- [11] Freund, Caroline and Emanuel Ornelas, 2010. “Regional Trade Agreements,” *Annual Review of Economics* 2, 139-167.
- [12] Grossman, Gene M. and Elhanan Helpman, 1995. “The Politics of Free-Trade Agreements.” *American Economic Review* 85, 667-690.
- [13] Horn, Henrik, Giovanni Maggi, and Robert W. Staiger, 2010. “Trade Agreements as Endogenously Incomplete Contracts.” *American Economic Review* 100(1), 394-419.
- [14] Kennan, John and Raymond Riezman, 1990. “Optimal Tariff Equilibria with Customs Unions?” *Canadian Journal of Economics* 23, 70-83.
- [15] Krishna, Pravin, 1998. “Regionalism and Multilateralism: A Political Economy Approach.” *Quarterly Journal of Economics* 113, 227-251.
- [16] Krishna, Pravin and Devashish Mitra, 2005. “Reciprocated Unilateralism in Trade Policy.” *Journal of International Economics* 65, 461–487.
- [17] Mrázová, Monika, David Vines, and Ben Zissimos, 2013. “Is the GATT/WTO’s Article XXIV Bad?” *Journal of International Economics* 89(1), 216-232.
- [18] Olarreaga, Marcelo, Isidro Soloaga, and Alan Winters, 1999. “What’s Behind MERCOSUR’s Common External Tariff?” Policy Research Working Paper Series 2231, World Bank.
- [19] Ornelas, Emanuel, 2005a. “Trade Creating Free Trade Areas and the Undermining of Multilateralism.” *European Economic Review* 49, 1717-1735.
- [20] Ornelas, Emanuel, 2005b. “Rent Destruction and the Political Viability of Free Trade Agreements.” *Quarterly Journal of Economics* 120, 1475-1506.
- [21] Saggi, Kamal and Halis M. Yildiz, 2010. “Bilateralism, Multilateralism, and the Quest for Global Free Trade.” *Journal of International Economics* 81, 26-37.
- [22] Saggi, Kamal, Alan Woodland and Halis M. Yildiz, 2013. “On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions.” *American Economic Journal: Microeconomics*, 5(1), 63-99.

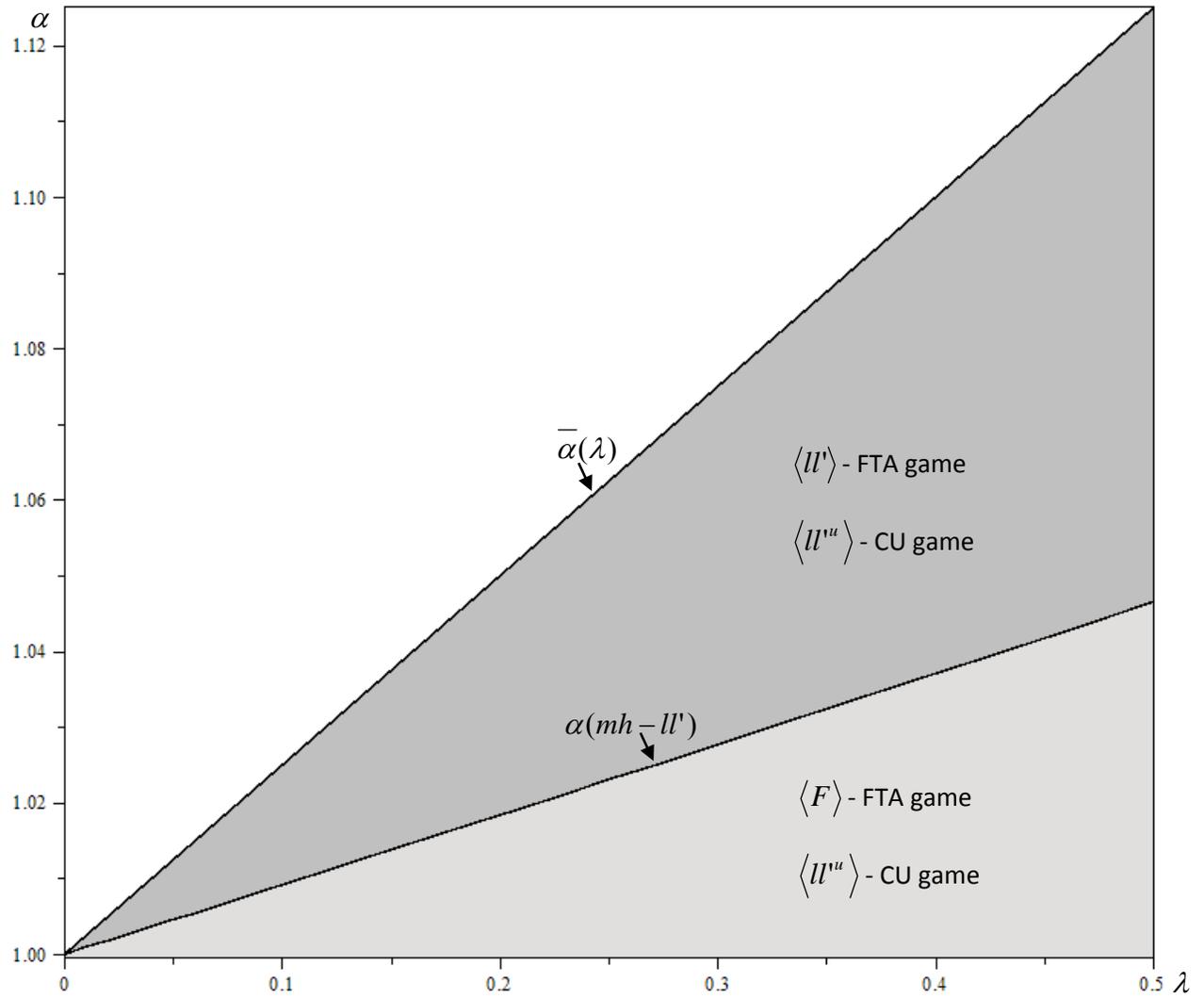


Figure 1: Equilibrium outcomes under asymmetry