On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions

Kamal Saggi*, Alan Woodland† and Halis Murat Yildiz‡

March 29, 2011

Abstract

This paper analyzes a game of trade policy (called *Bilateralism*) between three countries in which each country chooses whether to liberalize trade preferentially in the form of a Customs Union (CU), multilaterally, or not at all. We also analyze a restricted version of this game (called *Multilateralism*) under which countries do not have the option to form CUs. The analysis sheds light on the relationship between multilateral and preferential trade liberalization as sanctioned by GATT Article XXIV. We find that when countries have symmetric endowments, global free trade can be achieved without permitting CUs. Allowing for asymmetry, we isolate circumstances where Article XXIV helps further the cause of multilateral liberalization as well as when it does not. Furthermore, we show that Article XXIV's stipulation – that countries forming a CU not raise tariffs on outsiders – fails to make multilateral liberalization any more attractive to countries. However, such a tariff restriction does lower the adverse impact of a CU on the non-member.

Keywords: Customs Unions, preferential trade agreements, multilateral trade liberalization, GATT, WTO. *JEL Classifications*: F13, F12.

*Department of Economics, Vanderbilt University, Nashville, TN 37235-1828. Phone: 615-322-3237; e-mail: kamal.saggi@vanderbilt.edu. Parts of this paper were written during my visit to the Stanford Center for International Development (SCID), Stanford University, Palo Alto, CA. I am grateful to SCID researchers and its administrative staff for providing me with an excellent research environment.

†School of Economics, Australian School of Business, The University of New South Wales, NSW 2052, Australia. Phone: (61)-2-9385-9707; e-mail: A.Woodland@unsw.edu.au.

‡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

Article XXIV of the General Agreement on Tariffs and Trade (GATT) permits member countries of the World Trade Organization (WTO) to form preferential trade agreements (PTAs) such as free trade agreements (FTAs) and customs unions (CUs) under which PTA members can grant tariff reductions to each other that they do not extend to other WTO members. Empirical evidence indicates that WTO members have made rather liberal use of Article XXIV: as per the WTO's official web-site, as of Feb 2010, the WTO had received notification of 462 such arrangements, of which 345 were notified under Article XXIV.[1] While FTAs constitute an overwhelming majority of PTAs, the existing CUs involve some of the major economies of the world: for example, the Latin American CU MERCOSUR counts Argentina, Brazil, Paraguay, and Uruguay as its members while the EC (27) – a CU that extends across both goods and services – comprises of most major European economies. As a result, it is important to obtain a better understanding of the factors that give rise to CUs and the effect CUs have on the multilateral trading system.

The sanctioning of discriminatory trade agreements by GATT Article XXIV and the complicated web of global tariffs that has resulted from their pursuit by WTO member countries raises some uncomfortable questions about the very structure of GATT. Indeed, Article XXIV appears to be in direct conflict with the first and the most fundamental Article of GATT – i.e., the most favored nation (MFN) clause that forbids WTO members from pursuing discriminatory trade liberalization. Bhagwati (1991) has argued quite forcefully that PTAs are fundamentally incompatible with the WTO's stated goal of multilateral trade liberalization and many observers have wondered whether the multilateral trading system would function more effectively if the exception to non-discriminatory

[1] About 271 such agreements are already in force with the number expected to reach 400 during 2010. Mongolia is the only WTO member that does not participate in any PTA and most WTO members belong to multiple PTAs. Indeed, one even observes major PTAs in discussion with each other regarding mutual liberalization.

trade liberalization provided by Article XXIV were simply absent from GATT.

While Article XXIV sanctions PTAs, it does so only under certain conditions. In particular, Article XXIV requires that (1) a PTA should cover "substantially all trade" between members; (2) a PTA should result in significant trade liberalization among members almost to the point that it leads to free trade amongst them; and (3) that PTA members not raise tariffs on non-members – a condition that appears to be an attempt to safeguard the interests of those left outside such discriminatory trade agreements. Of course, these conditions do not necessarily imply that Article XXIV is successful in protecting the interests of non-members or that PTAs satisfying the requirements of Article XXIV are consistent with multilateral trade liberalization.

This paper attempts to isolate the consequences of Article XXIV for the process of multilateral trade liberalization. To accomplish this objective, we derive and compare the stable (or coalition proof) Nash equilibria of two games of trade liberalization between three countries with potentially asymmetric endowments. Under the first game (called *Bilateralism*), countries are free to liberalize trade bilaterally, multilaterally, or not at all. The bilateral option takes the form of a CU under which members eliminate tariffs on each other while imposing common external tariffs on the non-member. Under the second game, countries must either liberalize multilaterally or not at all. This restricted game – called *Multilateralism* – is an attempt to capture a world without Article XXIV of GATT. By comparing equilibrium outcomes under Bilateralism with those under Multilateralism, we isolate the consequences of the exception to non-discriminatory trade liberalization provided by GATT Article XXIV. The most attractive feature of our approach is that it endogenizes both the *nature* and the *degree* of trade liberalization – something that has been done rather infrequently in the rather vast literature on PTAs.[2]

Consistent with actual trade negotiations and the WTO's MFN principle, under our Multilateralism game a pair of countries undertaking reciprocal trade liberalization are

[2]See Bhagwati et. al. (1999) for a collection of some of the key papers in the area.

required to extend their respective tariff reductions to the third country. We denote such an *MFN-consistent trade agreement* between countries i and j by $\langle ij \rangle^m$. By contrast, under a discriminatory CU $\langle ij \rangle$ members i and j impose their optimal external tariffs on the non-member country. In our model, the tariff implemented by a country as a participant in the MFN-consistent agreement $\langle ij \rangle^m$ is lower than its optimal external tariff as a member of the CU $\langle ij \rangle$. As a result, the non-participating country (i.e., k) is worse off under the CU $\langle ij \rangle$ relative to the MFN-consistent agreement $\langle ij \rangle^m$ due to two separate reasons. First, it faces discriminatory tariffs in both export markets under $\langle ij \rangle$ whereas no such discrimination exists under $\langle ij \rangle^m$. Second, the external tariffs of countries i and j under $\langle ij \rangle$ are higher than those under $\langle ij \rangle^m$.

When countries are fully symmetric with respect to their endowments, we find that multilateral free trade emerges as the unique stable equilibrium under both Bilateralism and Multilateralism. In other words, under symmetry, global free trade obtains and a CU simply does not arise in equilibrium and the option to discriminate provided by Article XXIV is not exercised. This result suggests that heterogeneity across countries is likely to be a crucial determinant of why WTO members sometimes end up preferring discriminatory trading arrangements to non-discriminatory ones. Indeed, we show that when endowments are asymmetric across countries, the option to form CUs has important ramifications for the types of agreements that emerge in equilibrium. Furthermore, the nature of underlying asymmetry plays a rather subtle role in determining the relationship between bilateral and multilateral trade liberalization.

We first examine a scenario where one country (called country s) has a relatively smaller endowment than the other two (called l and l'). Under such a pattern of asymmetry, we find that the option to form CUs can actually increase the likelihood of achieving global free trade. The intuition for this result is as follows: opting out of global free trade is relatively costlier for a country when it faces a discriminatory CU between the other two countries – as it does under Bilateralism – than when it faces an MFN-consistent agree-

ment between them – as it does under Multilateralism. As a result, absent the threat of a discriminatory CU, the smaller country is less willing to liberalize multilaterally. However, if endowments are not too symmetric across countries, global free trade fails to arise under both Bilateralism and Multilateralism. If so, the equilibrium agreement reached under Multilateralism is welfare-superior to Bilateralism due to its non-discriminatory nature. An important practical implication of this result is that if global free trade is truly infeasible (i.e., cannot be reached under either Bilateralism or Multilateralism), then a purely multilateral approach to trade liberalization is preferable since it delivers a superior trade agreement.

When we consider a pattern of asymmetry where one country is relatively bigger (called l) than the other two (called s and s'), the argument in favor of Multilateralism becomes even stronger. First, even under this type of asymmetry, Multilateralism yields a superior trade agreement when free trade is infeasible under both games. Perhaps more interesting is the fact that, under this alternative pattern of asymmetry, global free trade arises over a larger parameter space under Multilateralism (i.e., when countries cannot form bilateral CUs). The intuition is that under Bilateralism each of the smaller countries has a joint incentive to exclude the larger country and to impose optimal external tariffs on its relatively large volume of exports. This result lends support to the view that discriminatory trade agreements can end up supplanting multilateral liberalization.

In the existing literature on PTAs, perhaps the two most closely related papers to ours are: Aghion et. al. (2007) and Saggi and Yildiz (2010).[3] While Aghion et. al. (2007) also provide a comparison of bilateral and multilateral liberalization, their approach is fundamentally different from ours. In particular, in their multilateral bargaining protocol countries must choose between global free trade and no agreement, whereas in our ap-

[3]In a recent paper, Seidmann (2009) develops a three-country bargaining model of trade negotiations in which bilateral agreements arise only if members gain more than outsiders and this alters the status quo in favor of PTA members during subsequent negotiations. Thus, PTAs can help improve the strategic position of members during future negotiations. Such a strategic incentive for PTA formation does not arise in our simultaneous game. Another important difference between his approach and ours is that, by design, PTAs cannot be conducive to the cause of global free trade in his model.

proach two countries can undertake reciprocal trade liberalization so long as each extends its respective tariff reduction to the third country on an MFN basis. This implies that in Aghion et al. (2007), by voting against free trade any country (say country k) can ensure that the status quo continues to prevail under multilateral bargaining. Under our approach to multilateral negotiations, a country that does not wish to liberalize ends up facing the MFN-consistent agreement $\langle ij \rangle^m$.[4] Thus, unlike in their model, the relative effects of the non-discriminatory agreement $\langle ij \rangle^m$ and the discriminatory CU $\langle ij \rangle$ play a crucial role in our analysis.

Our analysis of CUs complements and builds on that of Saggi and Yildiz (2010) who address related questions in the context of FTAs, the second major type of PTAs sanctioned by GATT Article XXIV.[5] The two types of PTAs differ in an important way: while FTA members impose individually optimal tariffs on non-members, CUs coordinate their external tariffs.[6] Naturally, this suggests that the potential adverse impact of a CU on non-members might be larger. This fear finds some support in the existing literature: in many models, the tariff complementarity effect – i.e., the tendency of PTA members to lower their tariffs on non-members – is generally weaker for a CU relative to an FTA. An important implication of this finding is that one of the key conditions of Article XXIV – i.e., PTA members should not raise tariffs on non-members – is more likely to bind for a CU relative to an FTA. In the model we present below, there exist circumstances where this indeed happens. To shed light on the potential effects of Article XXIV's constraint

[4] Our approach differs from Aghion et. al. (2007) in several other ways as well. First, they focus on FTAs while we consider CUs. Second, in our model, *all* countries are free to propose and negotiate agreements whereas in their model the follower countries must choose whether or not to join an agreement proposed by the leading country. Finally, unlike Aghion et. al. (2007), we do not allow transfers between different coalitions. This is important because they show that when transfers are possible and global free trade maximizes aggregate welfare, it emerges as the equilibrium under both sequential and multilateral bargaining. This is not the case in our model since transfers are not permitted.

[5] Several papers have considered the relationship between an endogenously determined FTA between two countries and its impact on incentives for multilateral liberalization – see, for example, Krishna (1998), and Ornelas (2005a and 2005b). Our contribution is to provide an equilibrium theory in which the choice between bilateral CUs and multilateral liberalization itself is endogenous.

[6] An important consequence of this difference between the two types of PTAs is that an FTA member is free to sign *another* FTA with an existing non-member *without* needing consent of an existing FTA partner whereas a CU member cannot do so.

on the external tariffs of CU members, we also derive equilibrium trade agreements under the assumption that the maximum tariff a CU member can charge equals its tariff prior to the formation of the CU (i.e., its tariff under the status quo). Our main finding here is that while this tariff restriction on CU members does not make the attainment of multilateral free trade any more likely, it does yield a constrained CU that is preferable from an aggregate welfare perspective.[7]

We next briefly discuss three strands of existing literature that overlap to some degree with our paper. The first related stream of literature comprises of models of repeated interaction that examine how the exogenous formation of a PTA affects incentives for self-enforcing multilateral tariff cooperation on the part of members and non-members – see Bagwell and Staiger (1997a, 1997b, and 1998), Bond and Syropoulos (1996), Hadjiyiannis (2004), and Saggi (2006).[8] A second closely related line of research examines the type of trading blocs that emerge as stable equilibria when the number and the sizes of trading blocs are endogenously determined – see Yi (1996), Goyal and Joshi (2006), and Furusawa and Konishi (2007).[9] Finally, the third related strand studies the endogenous formation of CUs using a cooperative game theory approach – see Riezman (1985), Kennan and Riezman (1990), Riezman (1999), and Melatos and Woodland (2007). While we depart from these lines of research in several ways, perhaps the most important difference is that we evaluate the relative pros and cons of preferential and multilateral trade liberalization in a model where the choice between the two types of liberalization is endogenous.

[7] Mrazova et al. (2010) examine the implications of the constraint Article XXIV imposes on a CU's external tariff in the oligopoly model of trade with an arbitrary number of countries and find that while this constraint makes CU formation less attractive, it can also lead to a reduction in world welfare when free trade fails to obtain.

[8] Unlike others, Bagwell and Staiger (1997a) also consider the period of transition during which a CU arises and find that while the early stages of CU formation facilitates multilateral liberalization, the harmony between CUs and multilateral trade liberalization is only temporary. Hadjiyiannis (2004) extends the Bagwell Staiger (1998) competing exporters model of trade to allow for asymmetric discount factors across countries. He examines a scenario where two countries are patient relative to the third and finds that the impact of a PTA depends on which countries form the PTA. Finally, Freund (2000) takes the multilateral tariff as exogenous and asks how it affects incentives for cooperation among PTA members.

[9] In a three country framework, Furusawa (2002) examines the choice between preferential arrangement and open regionalism when countries form a trading bloc.

2 Trade model

To endogenize the formation of trade agreements among asymmetric countries, we make use of an adapted version of the competing exporters model of Bagwell and Staiger (1997b and 1998). There are three countries: a, b, and c and three (non-numeraire) goods: A, B, and C. Each country's market is served by two competing exporters and I denotes the good that corresponds to the upper case value of i. For example, if $i = a$ then $I = A$. Country i is endowed with zero units of good I and e_i units of the other two goods.

The demand for good z in country i is as follows:[10]

$$d(p_i^z) = \alpha - p_i^z \text{ where } z = A, B, \text{ or } C \tag{1}$$

Since each country possesses only two goods while it demands all three, country i must import good I in order to consume it and it can import it from either trading partner. For example, country a imports good A from both countries b and c while it exports good B to country b and good C to country c.

Let t_{ij} be the tariff imposed by country i on its imports of good I from country j. Ruling out prohibitive tariffs yields the following no-arbitrage conditions for good I:

$$p_i^I = p_j^I + t_{ij} \ = p_k^I + t_{ik} \tag{2}$$

where $i, j, k = a, b, c$, and $i \neq j \neq k$. Let m_i^I be country i's imports of good I. Since country i has no endowment of good I, we have

$$m_i^I = d(p_i^I) = \alpha - p_i^I \tag{3}$$

Each country's exports of a good must equal its endowment of that good minus its local

[10] As is well known, the above demand functions can be derived from a utility function of the form $\sum_z u(c^z) + w$ where $u(\cdot)$ is quadratic; c^z denotes consumption of good z and $z = A, B$, or C; and w is the consumption of the numeraire good.

consumption:

$$x_j^I = e_j - [\alpha - p_j^I] \tag{4}$$

Market clearing for good I requires that country i's imports equal the total exports of the other two countries:

$$m_i^I = \sum_{j \neq i} x_j^I \tag{5}$$

Equations (2) through (5) imply that the equilibrium price of good I in country i equals:

$$p_i^I = \frac{1}{3}\left(3\alpha - \sum_{j \neq i} e_j + \sum_{j \neq i} t_{ij}\right) \tag{6}$$

Using these prices, the volume of trade is easily calculated. As is clear from equation (6), the price of good I in country i increases in its tariffs and decreases in the endowment levels of the other two countries. The effect of a country's tariff on its terms of trade is evident from equation (6): only a third of a given increase in either of its tariffs is passed on to domestic consumers with exactly two third of the tariff increase falling on the shoulders of foreign exporters.[11]

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

$$w_i = \sum_z CS_i^z + \sum_z PS_i^z + TR_i \tag{7}$$

Using equations (2) through (6) welfare of country i as a function of endowment levels and global tariffs can easily be calculated. Let aggregate world welfare be defined as the sum of each country's welfare

$$ww = \sum_i w_i \tag{8}$$

[11] By design the model examines country j's trade protection towards only good J (i.e. the only non-numeraire good that it imports).

We proceed as follows. First, we consider a three stage game of trade liberalization (which we refer to as *Bilateralism*) under which each country is free to pursue either (a) no trade liberalization or (b) bilateral trade liberalization or (c) multilateral trade liberalization.[12] After deriving Nash equilibria of this game and isolating those equilibria that are stable (more on this below), we next ask how equilibrium outcomes are affected if countries can choose only between options (a) and (c). This restricted game is called *Multilateralism*. Next, we compare the two games under the restriction that CU members cannot raise their tariffs on outsiders. These two experiments help shed light on the consequences of the exception to MFN provided under GATT Article XXIV as well as on the restriction imposed by it on the external tariffs of CU members.

3 Equilibrium under Bilateralism

The Bilateralism game proceeds as follows. In the first stage, each country announces the names of countries with whom it wants to form a customs union (CU). Country i's announcement is denoted by σ_i and its strategy set Ω_i consists of four possible announcements:

$$\Omega_i = \{\{\phi\}, \{j\}, \{k\}, \{M\}\} \tag{9}$$

where the announcement $\{\phi\}$ by country i is in favor of the status quo (or no trade liberalization); $\{j\}$ is in favor of a CU with only country j; $\{k\}$ is in favor of a CU with only country k; and $\{M\}$ is in favor of multilateral free trade. This announcement stage determines the global policy regime. Next, given the policy regime, countries impose their optimal external tariffs. Finally, given trade agreements and tariffs, international trade and consumption take place.

The Bilateralism game can yield the following policy regimes: (i) the status quo $\langle\Phi\rangle$

[12] Note that all countries have market power in the competing exporters model of Bagwell and Staiger (1997b and 1998) that we utilize. As a result, allowing for unilateral liberalization is not necessary – no country will choose to pursue it in this model.

prevails when no two announcements match or when everyone announces $\{\phi\}$; (ii) the CU $\langle ij \rangle$ is formed if countries i and j announce each other's name $\sigma_i = \{j\}$ and $\sigma_j = \{i\}$; (iii) free trade $\langle F \rangle$ obtains if $\sigma_i = \{M\}$ for all $i, j, k = a, b, c$.

One aspect of our Bilateralism game deserves comment. The CU $\langle ab \rangle$ obtains so long as country a and b call each other and neither of them calls country c, *regardless* of the nature of country c's announcement. This means that if $\sigma_a = \{b\}$ and $\sigma_b = \{a\}$, the excluded country (i.e., c) is indifferent between $\sigma_c = \{\phi\}$, $\sigma_c = \{a\}$, $\sigma_c = \{b\}$, or $\sigma_c = \{M\}$ since the outcome under any of these announcements on its part is the CU $\langle ab \rangle$. We settle such indifference on the part of the excluded country in favor of the most parsimonious announcement among the three – i.e., $\sigma_c = \{\phi\}$. This is reasonable since a proposal to form a CU is unlikely to be costless in the real world and a country facing no proposals from others would be better off not making any proposals of its own.

We next derive equilibrium trade agreements. Throughout the remainder of this section as well as in section 4, we maintain the following assumption:[13]

Assumption 1 (Symmetry):

$$e_i = e \text{ for all } i = a, b, c. \qquad \text{(symmetry)}$$

Let country i's welfare as a function of trade regime r be denoted by $w_i(r)$ where $r \in \{\langle \Phi \rangle, \langle ab \rangle, \langle ac \rangle, \langle bc \rangle, \text{ or } \langle F \rangle\}$. Also, let $\Delta w_i(r - v)$ denote the difference between country i's welfare under trade regimes r and v:

$$\Delta w_i(r - v) \equiv w_i(r) - w_i(v) \tag{10}$$

In accordance with Article I of GATT, we assume that under the status quo, each country imposes a non-discriminatory tariff on its trading partners: $t_{ij} = t_{ik} = t_i^{\phi}$ for all

[13] Calculations supporting the results reported in the rest of the paper are contained in the appendix.

$i, j, k = a, b, c$. Country i's optimal MFN tariff is easily calculated:

$$t_i^{\phi} \equiv Arg \max \; w_i(\Phi) = \frac{e}{4} \tag{11}$$

Under symmetry, $t_i^{\phi} = t^{\phi}$ for all $i = a, b, c$. If two countries form a CU, they remove tariffs on each other and impose jointly optimal external tariffs (denoted by t_i^u and t_j^u) on the non-member country.[14] The tariff pair (t_i^u, t_j^u) is chosen to solve:[15]

$$\max_{t_i^u,\, t_j^u} w_i(ij) + w_j(ij) \text{ subject to } t_{ij} = t_{ji} = 0 \tag{12}$$

Since each country is the unique importer of a good in our competing exporters model, the "market power effect" of a CU emphasized by Bagwell and Staiger (1997a) does not arise here since that effect arises only when CU members "compete" for imports.[16] As a result, the coordination of tariffs is beneficial to CU members only because each member internalizes the effect of its tariff on the export surplus of the other member. Furthermore, note that the constraint of zero internal tariffs does not bind for a CU since free internal trade is actually optimal for members.

When countries are symmetric, we have $t_i^u = t_j^u = t^u$ and the optimal external tariff of each CU member is given by

$$t^u = \frac{e}{5} \tag{13}$$

Note that, under symmetry, the formation of a CU induces each member country to *lower*

[14] Our simple formulation of a CU's tariff choice problem is intuitively appealing and in line with much of existing literature. However, Syropoulos (2003) has shown that the nature of the sharing rule of a CU with respect to tariff revenue can affect tariff preferences as well as the trade patterns of CU members in ways that can prevent the implementation of jointly optimal tariffs. An important insight of his analysis is that CU members have an incentive to influence their common tariffs not just for external terms-of-trade reasons but also for internal distributional purposes. Given the focus of our paper, we abstract from such considerations.

[15] The assumption that the CU maximizes the sum of national utilities is commonly employed in the literature. Issues of the delegation of tariff-setting authority and the choice of weights in the social welfare function are discussed by Gatsios and Karp (1991) and Melatos and Woodland (2007).

[16] In Bagwell and Staiger (1997a), countries forming a CU do not trade with each other at all and the CU is attractive to them only because it allows them to pool their market power and extract a larger terms of trade gain from non-members.

its tariff on the non-member relative to the status quo (i.e., the model exhibits tariff complementarity): $t^u < t^\phi$.[17] Furthermore, since tariffs are independent across countries, the non-member continues to impose its optimal MFN tariff on the member countries of the CU.

We are now ready to derive equilibrium trade agreements under Bilateralism. It is straightforward that the status quo 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. Which of the other two policy regimes – i.e., a CU $\langle ij \rangle$ and global free trade – are Nash equilibria? We address this next.

Denote the welfare of a country under a regime r with the relevant optimal tariffs substituted by $w^*(r)$. For example, when each country's tariff equals its optimal MFN tariff, the welfare of country i is given by $w_i^*(\Phi)$. Using calculations reported in the appendix, it is easy to show that a CU member has no unilateral incentive to defect from the agreement:

$$\Delta w_i^*(ij - \Phi) = \Delta w_j^*(ij - \Phi) > 0 \tag{14}$$

so that a CU is a Nash equilibrium. It is worth clarifying here that our model differs in an important way from the usual one-stage model of trade agreements where countries simply choose tariffs. In our model, the tariffs chosen by countries at the second stage depend upon the agreements reached in the first stage. In particular, if two countries agree to form a CU in the first stage, they essentially commit to choosing jointly optimal tariffs at the next stage. To see why this matters, suppose first stage announcements are given by $\sigma_a = \{b\}$, $\sigma_b = \{a\}$ and $\sigma_c = \{\phi\}$. Then, if country a unilaterally deviates and alters is announcement to $\sigma_a = \{\phi\}$ it does so with the understanding that at the tariff setting stage *all* countries will impose their optimal MFN tariffs on each other.

[17] Unlike Yi (1996), in our model the CU tariff does not violate Article XXIV under symmetry. It is noteworthy that tariff complementarity also arises in the general equilibrium model of Bond et. al. (2004). For empirical evidence regarding tariff complementarity in the context of the Latin American CU MERCOSUR, see Estevadeordal et. al. (2008).

It is noteworthy that the tariff complementarity effect in our model is not large enough to make the non-member better off under a CU relative to the status quo:

$$\Delta w_k^*(ij - \Phi) < 0 \tag{15}$$

Free trade is also a Nash equilibrium since no country has an incentive to unilaterally deviate from it:

$$\Delta w_k^*(F - ij) > 0 \tag{16}$$

To capture the process of CU formation in a more realistic fashion, we now refine the set of Nash equilibria by isolating those Nash equilibria that are *stable* or *coalition proof* – i.e., are immune to self-enforcing coalitional deviations as defined in Bernheim et. al. (1987).[18]

Which, if any, of the Nash equilibrium agreements described above are stable? For free trade to be stable, we need to rule out two joint deviations:

(JF1): Deviation of i and j from $\langle F \rangle$ to $\langle ij \rangle$.

(JF2): Deviation of i and j from $\langle F \rangle$ to $\langle \Phi \rangle$.

The first of these deviations - i.e., (JF1) – is easily dismissed since

$$\Delta w_i^*(F - ij) > 0 \tag{17}$$

Furthermore, it is immediate from (14) and (17) that joint deviation (JF2) will also not occur. In fact, $\langle \Phi \rangle$ is *not* a stable equilibrium because inequality (14) implies that the joint deviation of countries i and j from $\langle \Phi \rangle$ to $\langle ij \rangle$ is self-enforcing: any *further* deviation from $\langle ij \rangle$ on the part of either member reverts them back to $\langle \Phi \rangle$ under which both are worse-off. The CU $\langle ij \rangle$ also fails to be stable since all three countries have an incentive to jointly deviate from $\langle ij \rangle$ to $\langle F \rangle$ and this joint deviation is self-enforcing since each

[18] In our three-country model, a coalitional deviation is credible or self-enforcing if, taking the non-member's announcement as given, no member of the deviating coalition has an incentive to *further* deviate from its announcement as a part of the deviating coalition.

country's welfare under $\langle F \rangle$ is higher than that under the CU $\langle ij \rangle$ or the status quo.[19] Thus, we have:

Proposition 1: *Given symmetry, the status quo, a CU between two countries, and free trade are all Nash equilibrium trade agreements under Bilateralism. However, free trade is the only stable trade agreement.*

An important point to note is that when countries are symmetric, a CU between two countries is *not* a stable equilibrium. Under symmetry, the move from the status quo of optimal MFN tariffs to global free trade confers equal gains on all countries. As a result countries have an incentive to communicate with each other to attain the Pareto optimal outcome of global free trade. When countries make unilateral decisions and such communication is not possible, while free trade can arise as a Nash equilibrium, so can the status quo and a CU – outcomes that are decidedly worse from a global welfare perspective.

Proposition 1 raises the following question: can a bilateral CU arise as a stable equilibrium when countries are not symmetric in some respect? Before addressing this important question, we analyze the Multilateralism game under symmetry.

4 Equilibria under Multilateralism

Under the *Multilateralism* game, the strategy set of country i is $\Omega_i = \{\{\phi\}, \{M\}\}$, $j \neq k \neq i$ where $\{M\}$ is an announcement in favor of multilateral liberalization and $\{\phi\}$ an announcement against it. If all three countries announce in favor, they implement the jointly optimal set of tariffs which, in our model, are all equal to zero – i.e., a three-country agreement implies free trade. If only countries i and j announce in favor of multilateral liberalization $\{M\}$, they form the MFN-consistent agreement $\langle ij \rangle^m$ under which they jointly choose their optimal tariffs subject to the constraint that they *cannot*

[19] Note that under symmetry even the members of a CU are better off under free trade.

discriminate against country k.

Formally, countries i and j sign the agreement $\langle ij \rangle^m$ when individual country announcements are as follows: $\sigma_i = \{M\}$, $\sigma_j = \{M\}$, $\sigma_k = \{\phi\}$. Under the agreement $\langle ij \rangle^m$, countries i and j choose the tariff pair (t_i^m, t_j^m) to solve

$$\max_{(t_i^m,\, t_j^m)} w_i(ij^m) + w_j(ij^m) \text{ subject to } t_{ij} = t_{ik} = t_i^m \text{ and } t_{ji} = t_{jk} = t_j^m \tag{18}$$

Finally, if two (or more) countries do not announce $\{M\}$, the status quo $\langle \Phi \rangle$ prevails under which each country imposes its optimal MFN tariff on every other country.

As in the previous section, we maintain the assumption that countries are symmetric: $e_i = e$ for all $i = a, b, c$. Under symmetry, we must have $t_i^m = t_j^m = t^m$ and this jointly optimal MFN tariff is given by:

$$t^m = \frac{e}{7} \text{ where } t^m < t^\phi \tag{19}$$

Since $t^m < t^\phi$, it is immediate that countries that sign the MFN-consistent agreement $\langle ij \rangle^m$ lower their tariffs on each other as well as on the non-participating country (i.e., k). One aspect of the MFN-consistent agreement $\langle ij \rangle^m$ is noteworthy: since country k retains its optimal Nash tariff t^ϕ under $\langle ij \rangle^m$ while countries i and j cut tariffs on an MFN basis, it benefits from the trade liberalization undertaken by countries i and j without having to offer any liberalization in return.[20]

Next, note that $t^m < t^u$: i.e., country k faces higher tariffs in export markets when the other two countries implement a CU $\langle ij \rangle$ relative to when they sign the multilateral agreement $\langle ij \rangle^m$. To make matters worse, country k is also subject to *discriminatory treatment* in both its export markets under the CU $\langle ij \rangle$ – while countries i and j face zero

[20] In a competitive general equilibrium model, Raimondos-Møller and Woodland (2006) have shown that if coordinated non-discriminatory tariff reforms by a subset of countries are accompanied by appropriate income transfers between them, reforming members can make themselves strictly better off without having an adverse effect on non-members. In our model, while there are no transfers, the MFN-consistent agreement $\langle ij \rangle^m$ makes all countries strictly better off under symmetry. However, as we note below, the more relevant question here is how the non-member fares under $\langle ij \rangle^m$ relative to the CU $\langle ij \rangle$.

tariffs in each other's market under $\langle ij \rangle$, country k faces the tariff t^u. By contrast, such discriminatory treatment is absent under the agreement $\langle ij \rangle^m$. Therefore, the welfare of the non-member country (k) is strictly lower under the CU $\langle ij \rangle$ compared to the agreement $\langle ij \rangle^m$. Of course, as noted above, the flip side of this is that countries i and j are strictly better off under the CU $\langle ij \rangle$ relative to the multilateral agreement $\langle ij \rangle^m$ since only under the latter agreement do they have to abide by the non-discrimination constraint. These results are formalized in Section 5 as Lemma 1.

We now derive equilibrium agreements under Multilateralism. As under Bilateralism, the status quo $\langle \Phi \rangle$ is a Nash equilibrium under Multilateralism. In order to check whether $\langle ij \rangle^m$ is also Nash, we need to rule out two unilateral deviations:

(UM1): Deviation of i from $\langle ij \rangle^m$ to $\langle \Phi \rangle$

(UM2): Deviation of k from $\langle ij \rangle^m$ to $\langle F \rangle$.

A member country has no incentive to break the multilateral agreement $\langle ij \rangle^m$ since

$$\Delta w_i^*(ij^m - \Phi) > 0 \tag{20}$$

whereas the outside country (k) benefits from joining the agreement $\langle ij \rangle^m$ thereby converting it to $\langle F \rangle$:

$$\Delta w_k^*(F - ij^m) > 0 \tag{21}$$

Thus, under symmetry the multilateral agreement $\langle ij \rangle^m$ fails to be a Nash equilibrium because the outside country has an incentive to alter its stance in favor of the agreement thereby converting it to global free trade.[21] This result is interesting because it says that while the non-discriminatory trade liberalization that occurs under $\langle ij \rangle^m$ makes country k better off relative to the status quo, the extent of such liberalization is not large enough under symmetry to make country k opt out of the agreement.

The only remaining question is whether free trade is a Nash equilibrium. The answer

[21] We show later that such an agreement can indeed be a Nash equilibrium under asymmetry.

is in the affirmative: the only possible unilateral deviation that can occur from free trade is (UM2) and we have already argued that this deviation does not occur. Furthermore, it is clear that the status quo $\langle\Phi\rangle$ is not stable since all three countries benefit from deviating from $\langle\Phi\rangle$ to $\langle F\rangle$ from which there are no further unilateral or coalitional deviations – see inequalities (20) and (21). We thus have:

Proposition 2: *Given symmetry, the status quo and free trade are both Nash equilibrium trade agreements under Multilateralism. However, free trade is the only stable agreement.*

A comparison of Propositions 1 and 2 shows that when endowments are symmetric across countries, a multilateral approach to trade liberalization is sufficient to reach global free trade. One implication of this result is that if the gains from global trade liberalization were spread equally across countries (something that necessarily happens when countries have symmetric endowments), the freedom to pursue bilateral CUs is irrelevant for the ultimate objective of achieving global free trade. Given this result, the natural question is: do there exist circumstances under which the freedom to pursue bilateral CUs matters? We show next that such a possibility arises when endowments are sufficiently asymmetric across countries.

5 Liberalization among asymmetric countries

From hereon we drop Assumption 1 and begin by deriving optimal tariffs under asymmetry. If country i does not participate in any trade agreement, it chooses a non-discriminatory (or MFN) tariff t_i^ϕ to maximize its own welfare:

$$t_i^\phi \equiv Arg\max\ w_i(\Phi) = \frac{e_j + e_k}{8} \tag{22}$$

Note that country i's optimal tariff t_i^ϕ increases with the endowments of its trading partners. Similar to (13), when countries i and j form a CU $\langle ij \rangle$, they abolish tariffs on each other and choose their external tariffs to maximize their joint welfare $w_i(ij) + w_j(ij)$.[22] We have:

$$t_i^u = \frac{2e_k - e_j}{5} \text{ and } t_j^u = \frac{2e_k - e_i}{5} \tag{23}$$

It is easy to see that each CU member's external tariff increases with the endowment of the non-member whereas it decreases with that of its CU partner. This is intuitive. If the endowment of a CU member country increases, its exports to the partner country's market increase relative to the non-member and this induces the partner to reduce its tariff on the non-member.

As noted earlier, to minimize the potential harmful effects of PTAs on non-members, Article XXIV article requires that member countries not raise their external tariffs on non-members. Such a constraint on the tariff choices of country i as a CU member binds iff $t_i^u > t_i^\phi$. We have

$$t_i^u \geq t_i^\phi \Leftrightarrow 11e_k \geq 13e_j \tag{24}$$

For now, we ignore the restriction imposed by the Article XXIV and assume that a CU can impose its optimal tariffs. In Section 6, we examine the case where CUs must abide by this restriction on their external tariffs.

Under the multilateral agreement $\langle ij \rangle^m$ countries i and j choose the pair (t_i^m, t_j^m) to maximize $w_i(ij^m) + w_j(ij^m)$. We have

$$t_i^m = \frac{2e_k - e_j}{7} \text{ and } t_j^m = \frac{2e_k - e_i}{7} \tag{25}$$

As under symmetry, the comparison of the external tariffs under $\langle ij \rangle$ and $\langle ij \rangle^m$ implies that (i) the non-member faces lower tariffs under $\langle ij \rangle^m$ relative to $\langle ij \rangle$: $t_i^m < t_i^u$ and $t_j^m <$

[22] It is worth noting that under asymmetry it is still optimal for the CU member countries to reduce the internal tariffs to zero.

t_j^u; (ii) the non-member country is subject to *discriminatory treatment* in each member country's market under the CU $\langle ij \rangle$ relative to the agreement $\langle ij \rangle^m$. This implies:[23]

Lemma 1: *The welfare comparison of* $\langle ij \rangle$ *and* $\langle ij \rangle^m$ *from the perspective of members and the non-member is as follows:*

(i) The non-member is strictly worse off under $\langle ij \rangle$ *relative to* $\langle ij \rangle^m$*:* $w_k^*(ij) < w_k^*(ij^m)$*; and*

(ii) if $e_i = e_j$*, the member countries are strictly better off under* $\langle ij \rangle$ *relative to* $\langle ij \rangle^m$*:* $w_i^*(ij) = w_j^*(ij) > w_i^*(ij^m) = w_j^*(ij^m)$.

As explained before, the first part of the lemma is an immediate consequence of the higher external tariff under $\langle ij \rangle$ relative to $\langle ij \rangle^m$ and the discriminatory nature of the CU $\langle ij \rangle$. Symmetric member countries are strictly better off under the CU $\langle ij \rangle$ relative to the multilateral agreement $\langle ij \rangle^m$ since only under the latter agreement do they have to abide by the non-discrimination constraint.

An important aspect of our model deserves comment. Recall that the member countries of a CU import different goods from the non-member country since each country is a unique importer of a good. As a result, the formation of a CU requires member countries to alter their tariffs *solely* because they internalize the export surplus of the partner country and not due to any enhanced market power on their part since their tariffs apply to different goods. An implication of this is that the gains from the formation of a CU are not evenly split across countries when their endowments differ. The larger exporter in a CU (i.e., the one with the larger endowment) gains relatively more from the formation of the CU since the tariff it faces in the other member's market is relatively higher under the status quo.

Before proceeding further, we note that a restriction on the degree of endowment asymmetry must be imposed for the underlying pattern of trade assumed in the model (i.e., the competing exporters structure) to remain valid. To this end, let $x_j^I(r)$ denote

[23] Welfare levels under all possible regimes are reported in the appendix and these can be used to prove Lemmas 1-3.

the export of country j of good I to country i under regime any regime r. We show in the appendix that country j exports good I to country i under all trade regimes iff $3e_k \leq 5e_j$ or

$$x_j^I(r) \geq 0 \text{ iff } 3e_k \leq 5e_j \tag{26}$$

From hereon we assume that the above condition holds.

To derive equilibrium agreements, it is useful to briefly examine the incentives of asymmetric countries for bilateral and multilateral trade liberalization.[24] We begin by stating the following lemma:

Lemma 2: *Let country j be either a CU or free trade partner of country i under regime r where $r = \langle ij \rangle$ or $\langle F \rangle$ but not under regime v where $v = \langle \Phi \rangle$, $\langle ik \rangle$, or $\langle jk \rangle$. Then, the following hold:*
(i) $\frac{\partial \Delta w_i^*(r-v)}{\partial e_j} < 0 < \frac{\partial \Delta w_i^*(r-v)}{\partial e_i}$;
(ii) $\frac{\partial \Delta w_i^*(ij-\phi)}{\partial e_k} > 0$ and $\frac{\partial \Delta w_i^*(F-ij)}{\partial e_j} > 0$.

The intuition for part (i) of Lemma 2 is as follows. The smaller a country's endowment, the lower its volume of exports and the less it benefits from tariff reductions granted by others. Similarly, the larger the endowment of a country's partner, the larger the terms of trade loss suffered by it from granting the partner free access to its market. Thus, a country's willingness to eliminate its tariff on a trading partner depends positively on its own endowment and negatively on that of its partner. The second part of the lemma captures two related points. The inequality $\frac{\partial \Delta w_i^*(ij-\phi)}{\partial e_k} > 0$ says that, beginning at the status quo $\langle \Phi \rangle$, the advantage gained by country i in country j's market from signing the CU $\langle ij \rangle$ increases in country k's size, making such a CU more valuable from its welfare perspective. Recall that both countries i and k export the same good to country j (i.e.,

[24] In the competing exporters model of Bagwell and Staiger (1997b and 1998) utilized by us, asymmetry in endowments translates directly into asymmetries of export volumes: an increase in a country's endowment increases its exports of non-numeraire/protected goods without affecting its imports of such goods since the model lacks any income effects. Indeed, the country with the largest endowment of non-numeraire goods faces the smallest imports of such goods.

they are competing exporters). Thus, a country always prefers to form a CU with the smaller of its two trading partners:

$$w_i^*(ij) \geq w_i^*(ik) \text{ iff } e_k \geq e_j \tag{27}$$

The inequality $\frac{\partial \Delta w_i^*(F-ij)}{\partial e_j} > 0$ says that when country j is already a CU partner of country i, country i's welfare gain from moving to multilateral free trade increases with the endowment of country j. This is because the value of the preferential treatment enjoyed by country i in country j's market under $\langle ij \rangle$ decreases with the endowment of the competing exporter j.

How do the incentives of countries to form (or join) an MFN-consistent trade agreement depend on the underlying endowment structure?

Lemma 3: *Under Multilateralism, the following hold:*
(i) $\frac{\partial \Delta w_i^*(ij^m-\phi)}{\partial e_i} > 0$, $\frac{\partial \Delta w_i^*(ij^m-\phi)}{\partial e_j} < 0$ *and* $\frac{\partial \Delta w_i^*(ij^m-\phi)}{\partial e_k} < 0$*; and*
(ii) $\frac{\partial \Delta w_i^*(F-ij^m)}{\partial e_i} > 0$, $\frac{\partial \Delta w_i^*(F-ij^m)}{\partial e_j} < 0$ *and* $\frac{\partial \Delta w_i^*(F-ij^m)}{\partial e_k} < 0$.

The intuition underlying the inequalities reported in Lemma 3 is quite analogous to that which underlies the results reported in Lemma 2 with one exception – i.e., whereas $\frac{\partial \Delta w_i^*(ij-\phi)}{\partial e_k} > 0$ under Bilateralism, the opposite is true under Multilateralism – i.e., $\frac{\partial \Delta w_i^*(ij^m-\phi)}{\partial e_k} < 0$. To see why this is the case recall that under the agreement $\langle ij \rangle^m$, countries i and j lower their tariffs on each other as well as on country k whereas under the CU $\langle ij \rangle$ they only lower tariffs on each other. As a result, the larger is country k's endowment, the smaller the increase in the export surplus that countries i and j obtain from the agreement $\langle ij \rangle^m$ since their rival exporter (i.e., country k) captures a larger share of their markets.

Part (i) of Lemma 2 implies that a country prefers to enter into an MFN-consistent

agreement with the smaller of its two trading partners:

$$w_i^*(ij^m) \geq w_i^*(ik^m) \text{ iff } e_k \geq e_j \tag{28}$$

To highlight the crucial role played by asymmetry in our model, it proves instructive to focus on two special cases: (i) two countries (denoted by l and l') have larger endowments than the third (denoted by s) and (ii) two countries (denoted by s and s') have smaller endowments than the third (denoted by l). We consider each in turn.

6 If one country is smaller

Let the pattern of endowment asymmetry be given by:

Assumption 2a:

$$e_s = \frac{e}{\theta} < e_l = e_{l'} = e \text{ and } 1 \leq \theta \leq \frac{5}{3} \tag{29}$$

The restriction $\theta \leq \frac{5}{3}$ is necessary to ensure that the pattern of trade assumed by the competing exporters structure remains valid under asymmetry.

6.1 Equilibrium agreements

To avoid redundancy, we focus directly on stable agreements under Bilateralism (i.e., we skip the discussion of Nash equilibria). First consider the perspective of the larger countries. We know from (14) and (17) that starting from global free trade, under symmetry, two countries have no incentive to jointly deviate from $\langle F \rangle$ to $\langle \Phi \rangle$. Lemma 2 implies that $\frac{\partial w_l^*(F-\phi)}{\partial \theta} = \frac{\partial \Delta w_{l'}^*(F-\phi)}{\partial \theta} > 0$ and thus it follows that the larger countries have no incentive to deviate from $\langle F \rangle$ to $\langle \Phi \rangle$:

$$\Delta w_l^*(F - \Phi) = \Delta w_{l'}^*(F - \Phi) > 0 \text{ for all } \theta \tag{30}$$

Similarly, we know from (16) that under symmetry ($\theta = 1$), a country has no incentive to unilaterally deviate from free trade to become an outsider facing a CU. From Lemma 2, we have $\frac{\partial \Delta w_l^*(F-sl')}{\partial \theta} = \frac{\partial \Delta w_{l'}^*(F-sl)}{\partial \theta} > 0$. Thus, it follows that

$$\Delta w_l^*(F - sl') = \Delta w_{l'}^*(F - sl) > 0 \text{ for all } \theta \tag{31}$$

Do the larger countries have an incentive to exclude the small country by deviating from $\langle F \rangle$ to $\langle ll' \rangle$? We know from (17) that this is not the case under symmetry ($\theta = 1$). Further, note from Lemma 2 that $\frac{\partial \Delta w_l^*(F-ll')}{\partial \theta} = \frac{\partial \Delta w_l^*(F-ll')}{\partial \theta} > 0$. This implies that

$$\Delta w_l^*(F - ll') = \Delta w_{l'}^*(F - ll') > 0 \text{ for all } \theta \tag{32}$$

Only two possible defections from free trade remain to be considered:

(UF): Unilateral deviation of country s from $\langle F \rangle$ to $\langle ll' \rangle$.

(JF): Joint deviation of countries s and l (or l') from $\langle F \rangle$ to $\langle sl \rangle$ (or $\langle sl' \rangle$).

Let $\theta_i(r - v)$ denote the critical degree of endowment asymmetry at which country i is indifferent between regimes r and v. Direct calculations show that the deviation (UF) occurs when country s is sufficiently smaller than others:

$$\Delta w_s^*(F - ll') \leq 0 \text{ iff } \theta \geq \theta_s(F - ll') \tag{33}$$

Now consider deviation (JF). It turns out that a larger country (say l) has an incentive to jointly deviate with the smaller country provided it is sufficiently small:

$$\Delta w_l^*(F - sl) \leq 0 \text{ iff } \theta \geq \theta_l(F - sl) \tag{34}$$

Similarly, we have

$$\Delta w_s^*(F - sl) \leq 0 \text{ iff } \theta \geq \theta_s(F - sl) \tag{35}$$

We show in the appendix that $\theta_l(F - sl) < \theta_s(F - sl)$ so that deviation (JF) occurs when $\theta \geq \theta_s(F - sl)$. Note further that this joint deviation is self-enforcing only when $\theta \leq \theta_s(sl - \Phi)$; otherwise, the smaller country has an incentive to further deviate unilaterally from $\langle sl \rangle$ to $\langle \Phi \rangle$:

$$\Delta w_s^*(sl - \Phi) \leq 0 \text{ iff } \theta \geq \theta_s(sl - \Phi) \tag{36}$$

The deviation conditions specified in (33), (34), (35) and (36) imply that *free trade is stable only when* $\theta \leq \theta_s(F - sl)$. By considering the stability of each of the other remaining Nash equilibria, we prove the following result in the appendix:

Proposition 3: *Given Assumption 2a, the stable equilibria under Bilateralism are as follows:*
(i) free trade obtains over $\theta \leq \theta_s(F - sl)$*;*
(ii) the CU $\langle sl \rangle$ *obtains over* $\theta_l(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$*; and*
(iii) the CU $\langle ll' \rangle$ *obtains over* $\theta \geq \theta_s(sl - \Phi)$*;*

– Figure 1 here –

Proposition 3 relates the degree of underlying asymmetry to the nature of stable agreements. Part (i) simply says that if the degree of endowment asymmetry is sufficiently small, free trade is uniquely stable. Part (ii) says that if the degree of endowment asymmetry is moderate, a CU between two asymmetric partners is stable.[25] When the degree of endowment asymmetry is sufficiently large, only a CU between the two larger countries is stable. Over this range, the smaller country faces relatively low tariffs under the CU $\langle ll' \rangle$ and it prefers to remain an outsider.

[25]Note that when $\theta_l(F - sl) \leq \theta \leq \theta_s(F - sl)$, both $\langle \{F\} \rangle$ and $\langle \{sl\} \rangle$ or $\langle \{sl'\} \rangle$ are stable. Since theory offers no guidance about which of these equilibria might be observed, we examine both of these possibilities hereafter.

6.2 Article XXIV and multilateral liberalization

In order to evaluate the effects of Article XXIV, we now analyze our Multilateralism game between asymmetric countries. Using arguments similar to those employed under symmetry, it can be shown that there are four possible Nash equilibria under Multilateralism: $\langle\Phi\rangle$, $\langle sl\rangle^m$, $\langle ll'\rangle^m$ and $\langle F\rangle$. From (20) and Lemma 3, it follows that the joint deviation of the two larger countries from $\langle\Phi\rangle$ to $\langle ll'\rangle^m$ is self-enforcing. Thus, $\langle\Phi\rangle$ is not stable. Similarly, inequality (21) and Lemma 3 imply that one of the larger countries (say l') has an incentive to unilaterally deviate from $\langle sl\rangle^m$ to $\langle\{F\}\rangle$ so that $\langle sl\rangle^m$ is not a stable agreement.

To see when and why the other two agreements are stable, first note that (30) implies that there can be no deviations from $\langle\{F\}\rangle$ to $\langle\Phi\rangle$. Furthermore, as argued above, inequality (21) and Lemma 3 imply that the larger country l' has no incentive to unilaterally deviate from $\langle\{F\}\rangle$ to $\langle sl\rangle^m$. In fact, the only deviation from free trade that we need to consider is the unilateral deviation of the smaller country to $\langle ll'\rangle^m$. It turns out this deviation occurs only when the degree of endowment asymmetry is sufficiently large:

$$\Delta w_s^*(F - ll'^m) \leq 0 \text{ iff } \theta \geq \theta_s(F - ll'^m) \tag{37}$$

It immediately follows that free trade is stable under Multilateralism when $\theta \leq \theta_s(F - ll'^m)$.

What if $\theta > \theta_s(F - ll'^m)$? We know from inequality (20) that $\Delta w_l^*(ll'^m - \Phi) > 0$ under symmetry ($\theta = 1$). Since $\frac{\partial \Delta w_l^*(ll'^m - \phi)}{\partial \theta} > 0$ (Lemma 3) we have

$$\Delta w_l^*(ll'^m - \Phi) > 0 \text{ for all } \theta \tag{38}$$

Then using inequalities (37) and (38) we can argue that the multilateral agreement $\langle ll'\rangle^m$ is stable over the range $\theta > \theta_s(F - ll'^m)$.

Proposition 4: *Given Assumption 2a, free trade is a stable equilibrium when* $\theta \leq \theta_s(F - ll'^m)$. *Otherwise, the MFN-consistent agreement* $\langle ll' \rangle^m$ *is stable.*

Figure 2 shows stable agreements under Multilateralism.

– Figure 2 here –

Recall that under Bilateralism, global free trade is stable when $\theta < \theta_s(F-sl)$ whereas it is stable under Multilateralism only when $\theta \leq \theta_s(F-ll'^m)$. A straightforward comparison of these critical thresholds along with Figures 1 and 2 delivers a central result of our paper:

Proposition 5: *Given Assumption 2a, the following hold:*
(i) $\theta_s(F - ll'^m) < \theta_s(F - sl)$; *and*
(ii) over the parameter range $\theta_s(F-ll'^m) < \theta \leq \theta_l(F-sl)$ *the unique stable agreement under Bilateralism is free trade whereas under Multilateralism it is the MFN-consistent agreement* $\langle ll' \rangle^m$.

Part (i) of proposition 5 says that free trade is stable over a *larger* parameter space when countries are free to form CUs relative to when they cannot. Part (ii) isolates the parameter range over which the freedom to pursue CUs is *necessary* for achieving global free trade. Over this range, because $\theta > \theta_s(F-ll'^m)$ the smaller country has an incentive to unilaterally deviate from free trade to $\langle ll' \rangle^m$ under Multilateralism whereas there exist no such incentive (either unilateral or coalitional) under Bilateralism. It is important to note that, under Multilateralism, the stability of free trade hinges on the *unilateral* deviation incentive of the smaller country whereas under Bilateralism it depends on the *coalitional* deviation incentive of the smaller country and one of the larger countries. As indicated earlier, the smaller country faces a higher tariff under a CU $\langle ll' \rangle$ relative to the agreement $\langle ll' \rangle^m$. Moreover, under the CU $\langle ll' \rangle$ the smaller country is discriminated against in its export markets whereas it suffers no such disadvantage under the agreement $\langle ll' \rangle^m$. The

joint effect of these two factors – a higher tariff and the discriminatory nature of the CU – makes opting out of global free trade relatively costlier for the smaller country.[26]

– Figure 3 here –

What if global free trade fails to obtain under both Bilateralism and Multilateralism? Under such a situation, one possible way forward is to ask how global welfare compares under the equilibrium agreements that obtain under these two games. From Proposition 3 we know that, when free trade is not feasible under Bilateralism – i.e., when $\theta > \theta_s(F-sl)$ – either $\langle sl \rangle$ or $\langle ll' \rangle$ are stable agreements. Similarly, Proposition 4 says that when global free trade does not obtain under Multilateralism, $\langle ll' \rangle^m$ emerges as the unique stable equilibrium.

To compare Bilateralism and Multilateralism when free trade fails to materialize, first consider the range $\theta_s(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$. Over this range, the CU $\langle sl \rangle$ is stable under Bilateralism while $\langle ll' \rangle^m$ obtains under Multilateralism. Compare global welfare under these agreements and using the formulae reported in the appendix we find that Multilateralism yields a superior outcome:

$$\Delta ww^*(ll'^m - sl) > 0 \tag{39}$$

The intuition for this result is straightforward: unlike the multilateral agreement $\langle ll' \rangle^m$, the CU $\langle sl \rangle$ diverts trade from country l' to country s and this problematic because, under non-discriminatory tariffs, country l' has a lower relative price of the good exported to country l.

[26]It is also important to note that Proposition 5 focuses only on the parameter range where free trade is uniquely stable under Bilateralism. However, we know from Proposition 3 that both free trade and an asymmetric CU are stable over $\theta_l(F - sl) \leq \theta \leq \theta_s(F - sl)$. Thus, one can argue that when $\theta_l(F - sl) \leq \theta \leq \theta_s(F - sl)$, the freedom to pursue CUs acts as a force in favor of trade liberalization in a weaker sense.

Similarly, when $\theta \geq \theta_s(sl - \Phi)$, the CU $\langle ll' \rangle$ arises under Bilateralism while the agreement $\langle ll' \rangle^m$ obtains under Multilateralism which is once again a superior outcome:

$$\Delta ww^*(ll'^m - ll') > 0 \tag{40}$$

Thus, when free trade is infeasible, the option to pursue CUs leads to inferior outcomes. What are the policy implications of this result? As any observer of the global policy landscape can see, international trade barriers are very much present today, have been with us essentially forever, and might always be part of the global policy landscape. If global free trade is truly out of reach, then there is an argument to be made in favor of excluding CUs from Article XXIV of GATT since non-discriminatory multilateral liberalization avoids the type of distortions that result from the formation of CUs.

Of course, aggregate world welfare does not necessarily speak to the fate of individual countries. In this regard, we can state the following:

Proposition 6: *Suppose Assumption 2a holds. Then, the following hold:*
(i) when $\theta_s(F - ll'^m) < \theta \leq \theta_s(F - sl)$, *the two larger countries are better off under Bilateralism whereas the smaller country is better off under Multilateralism.*
(ii) when $\theta_s(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$, *the smaller country and the larger country that is excluded from the CU (that arises under Bilateralism) are worse off under Bilateralism whereas the other larger country is better off under Multilateralism.*
(iii) when $\theta > \theta_s(sl - \Phi)$, *the two larger countries are better off under Bilateralism whereas the smaller country is better off under Multilateralism .*

Since world welfare is higher under free trade relative to $\langle ll' \rangle^m$, the first part of the proposition is a direct implication of inequality (37). Next, consider the second part. When $\theta_s(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$, suppose the CU $\langle sl \rangle$ is the stable agreement under Bilateralism while $\langle ll' \rangle^m$ obtains under Multilateralism. Using Lemma 1 and (27), we note $w_l^*(sl) > w_l^*(ll') > w_l^*(ll'^m)$. This implies that the larger country that ends up as a CU member (i.e., country l) is better off under Bilateralism relative to Multilateralism. The flip side of this is that the excluded larger country under Bilateralism prefers Multilateralism: $w_{l'}^*(sl) < w_{l'}^*(ll'^m)$. But why is the smaller country worse off under Bilateralism even

though it is a willing member of the CU $\langle sl \rangle$? Direct calculations establish that when $\theta_s(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$ we have $w_s^*(sl) < w_s^*(ll'^m)$. In other words, over this range, the smaller country would prefer to be an outsider facing an MFN-consistent agreement between the two larger countries rather than being a member of a CU, which is what ends up happening under Bilateralism.

Finally, the last part of the proposition is a direct implication of Lemma 1 since (i) the smaller country faces discriminatory treatment and higher external tariffs under $\langle ll' \rangle$ relative to $\langle ll' \rangle^m$; and (ii) the large countries, provided they are symmetric, are better off under CU $\langle ll' \rangle$ relative to the agreement $\langle ll' \rangle^m$ since only under the latter agreement do they have to abide by the non-discrimination constraint.

The combination of proposition 4 and proposition 5 implies that, when the freedom to pursue bilateral CUs is *necessary* for achieving global free trade, Bilateralism lead to higher world welfare. However, when free trade is not feasible under either regime, the smaller country's preference for Multilateralism is consistent with the world welfare ranking.

We now consider the case where one country is larger than the other two.

7 If one country is larger

Let the pattern of endowments be given by:

Assumption 2b: $e_s = e_{s'} = \frac{e}{\theta} < e_l = e$ and $\frac{5}{3} \geq \theta > 1$.

In other words, the endowments of countries s and s' are smaller than that of country l.

7.1 Equilibrium agreements

To conserve space, we directly state the following result and prove it in the appendix:

Lemma 4 *Suppose Assumption 2b holds. Then, there exist no unilateral or coalitional deviation from free trade that involves the large country.*

Using Lemma 4 and arguments analogous to those used in section 6, we can show:

Proposition 7: *Suppose Assumption 2b holds. Then, free trade is stable under Bilateralism if* $\theta \leq \theta_s(F - ss')$. *Otherwise, the CU* $\langle ss' \rangle$ *is stable.*

– Figure 4 here –

The above proposition hinges on the fact that the joint incentive of the two smaller countries to discriminate against the larger country determines the stability of free trade. As before, in order to examine how the ability to form CUs matters, consider the equilibrium outcomes of the Multilateralism game. This game has four Nash equilibria: $\langle \Phi \rangle$, $\langle ss' \rangle^m$, $\langle sl \rangle^m$ and $\langle F \rangle$. From (21) and Lemma 3 it follows that the larger country has no incentive to unilaterally deviate from $\langle F \rangle$ to $\langle ss' \rangle^m$:

$$\Delta w_l^*(F - ss'^m) > 0 \text{ for all } \theta \tag{41}$$

Thus, $\langle ss' \rangle^m$ *is not stable*. Note also that $\langle \Phi \rangle$ *is not stable* since the smaller countries have a joint incentive to deviate from $\langle \Phi \rangle$ to $\langle ss' \rangle^m$ and it is a self-enforcing deviation since

$$\Delta w_s^*(ss'^m - \Phi) > 0 \text{ for all } \theta \tag{42}$$

On the other hand, country s' has an incentive to unilaterally deviate from $\langle F \rangle$ to $\langle sl \rangle^m$ if θ is sufficiently large:

$$\Delta w_{s'}^*(F - sl^m) < 0 \text{ iff } \theta > \theta_{s'}(F - sl^m)$$

The smaller countries have incentives to jointly deviate from $\langle F \rangle$ to $\langle \Phi \rangle$ provided they

are sufficiently small relative to the larger country:

$$\Delta w_s^*(F-\Phi) < 0 \text{ iff } \theta > \theta_s(F-\Phi)$$

However, since $\theta_s(F-\Phi) > \theta_{s'}(F-sl^m)$, it is not a binding deviation for the stability of free trade. Thus $\langle F \rangle$ is stable iff $\theta \leq \theta_{s'}(F-sl^m)$ holds. Finally, country s has an incentive to unilaterally deviate from $\langle sl \rangle^m$ to $\langle \Phi \rangle$ if θ is sufficiently large:

$$\Delta w_s^*(sl^m-\Phi) < 0 \text{ iff } \theta > \theta_s(sl^m-\Phi)$$

Thus $\langle sl \rangle^m$ is stable iff $\theta_{s'}(F-sl^m) < \theta \leq \theta_s(sl^m-\Phi)$ holds. Therefore, we have the following result:

Proposition 8: *Suppose Assumption 2b holds. Then, the stable agreements under Multilateralism are as follows:*
(i) free trade obtains over $\theta < \theta_{s'}(F-sl^m)$; and
(ii) the MFN-consistent agreement $\langle sl \rangle^m$ obtains over $\theta_{s'}(F-sl^m) < \theta \leq \theta_s(sl^m-\Phi)$.

– Figure 5 here –

The above proposition shows that under Multilateralism the unilateral deviation incentive of one of the smaller countries determines the stability of free trade. Another important observation is that the set of stable equilibria is now empty when the larger country has a sufficiently big endowment relative to the smaller countries - when $\theta > \theta_s(sl^m-\Phi)$.

7.2 When a CU supplants multilateral free trade

What light does our model shed on the effects of Article XXIV when two countries are relatively smaller? We directly state our main result:

Proposition 9: *Suppose Assumption 2b holds. Then, when $\theta_s(F-ss') < \theta \leq \theta_{s'}(F-sl^m)$ holds, Multilateralism yields free trade as the stable equilibrium whereas Bilateralism yields the CU $\langle ss' \rangle$.*

– Figure 6 here –

Proposition 9 argues that, unlike the case of two larger and one smaller countries, free trade is stable over a *larger* parameter space when countries *cannot* form bilateral CUs. Thus over the parameter range $\theta_s(F-ss') < \theta \leq \theta_{s'}(F-sl^m)$, the freedom to pursue CUs prevents the attainment of global free trade. This is in sharp contrast to the results of Saggi and Yildiz (2010) who analyzed the case of FTAs. Together, these results suggest that the concern that PTAs might undermine the multilateral trading system maybe more relevant for CUs than for FTAs.

When countries are free to pursue CUs and $\theta_s(F-ss') < \theta \leq \theta_{s'}(F-sl^m)$, the smaller countries have a joint incentive to exclude the larger country via deviating from $\langle F \rangle$ to $\langle ss' \rangle$ and this deviation is self-enforcing since neither of the smaller countries has a unilateral incentive to further deviate. On the other hand, over the same range, there exists no unilateral or joint deviation from free trade under Multilateralism.

Why does the freedom to pursue CUs supplant free trade when the pattern of asymmetry is given by Assumption 2b (one country is smaller) whereas it has the opposite effect under Assumption 2a (one country is larger)? The intuition behind this result can be built in steps. First, note that regardless of the pattern of asymmetry, under Multilateralism it is the *unilateral deviation incentive* of a smaller country that determines the stability of free trade. Second, the benefit of the external tariff reduction enjoyed by the smaller country under the agreement $\langle ll' \rangle^m$ – relevant under Assumption 2a – is larger than that under $\langle sl \rangle^m$ – applicable under Assumption 2b. This in turn implies that when the pattern of asymmetry is given by assumption 2a, the smaller country is less willing to participate in free trade relative to the case when it is given by Assumption 2b – i.e., $\theta_{s'}(F-sl^m) > \theta_s(F-ll'^m)$.

The third point to note is that in the Bilateralism game, it is the *joint deviation incentive* of two countries to exclude the larger country from free trade – of countries s and l under Assumption 2a and countries s and s' under Assumption 2b – that determines

the stability of free trade. In this regard, we find that this joint incentive to exclude is stronger under Assumption 2b (when two countries are smaller than the third) relative to that under Assumption 2a because, in the former scenario, each of them benefits relatively more from being able to impose its optimal CU tariff on the relatively large volume of imports supplied by the non-member. This implies that the region over which free trade is stable under Bilateralism is larger under Assumption 2a: $\theta_s(F - ss') < \theta_s(F - sl)$.

These three points together help explain how the underlying pattern of endowment asymmetry ends up determining the role of Article XXIV. Relative to Assumption 2a, the pattern of asymmetry under 2b yields free trade over a *larger* parameter space in the Multilateralism game while it yields free trade over a *smaller* parameter space under the Bilateralism game. This is why Article XXIV ends up harming the prospects for global free trade under Assumption 2b whereas it has the opposite effect under Assumption 2a.

Finally, what if $\theta > \theta_{s'}(F - sl^m)$ and global free trade fails to obtain under both Bilateralism and Multilateralism? As before, under such a situation, we can ask how global welfare compares under the equilibrium agreements that obtain under the two games. We know that when $\theta > \theta_{s'}(F - sl^m)$, the agreement $\langle sl \rangle^m$ is stable under Multilateralism if $\theta \leq \theta_s(sl^m - \Phi)$ while the CU $\langle ss' \rangle$ is stable under Bilateralism. Comparing global welfare under these agreements yields

$$\Delta ww^*(sl^m - ss') > 0 \tag{43}$$

Once again, Multilateralism yields the more preferred outcome from an aggregate welfare perspective. The preferences of individual countries are as follows:

Proposition 10: *Suppose Assumption 2b holds. Then, the larger country is better off under Multilateralism relative to Bilateralism whereas the preference of each of the smaller countries is the opposite.*

When $\theta_s(F - ss') < \theta \leq \theta_{s'}(F - sl^m)$, the joint deviation of the smaller countries determines the stability of free trade and they are better off under the CU $\langle ss' \rangle$ relative

to free trade. Since world welfare is higher under free trade relative to the CU $\langle ss' \rangle$, this automatically implies that the larger country is better off under Multilateralism relative to Bilateralism over this parameter range. When $\theta > \theta_{s'}(F - sl^m)$, recall that $\langle sl \rangle^m$ is stable under Multilateralism if $\theta \leq \theta_s(sl^m - \Phi)$ while the CU $\langle ss' \rangle$ is stable under Bilateralism. Over this range it is immediate from Lemma 1 and (28) that $w_s^*(ss') > w_s^*(ss'^m) > w_s^*(sl^m)$, thus country s is better off under Bilateralism relative to Multilateralism. When $\theta > \theta_{s'}(F - sl^m)$, the same result obtains for country s': $w_{s'}^*(ss') > w_{s'}^*(sl^m)$ while the opposite result obtains for the larger country which is better off under Multilateralism relative to Bilateralism.

Our final task in this paper is to examine the consequences of the tariff restriction imposed on CU members by Article XXIV.

8 Constraining external tariffs of a CU

If CU members are prohibited from raising tariffs on non-members, none of our results under symmetry are affected since CU members actually find it optimal to *lower* their tariffs on non-members when countries have symmetric endowments.

Suppose endowments are as specified in Assumption 2a. Under this assumption, the only agreement for which the tariff constraint of Article XXIV can potentially bind is the CU between a larger country and the smaller country, say $\langle sl \rangle$. Specifically, under the CU $\langle sl \rangle$, the tariff restriction specified in Article XXIV binds for country l when $\theta > \overline{\theta} \equiv \frac{13}{11}$ since over this range $t_l^\phi < t_l^u$. It is easy to see that our entire analysis under Assumption 2a remains unchanged when $\theta \leq \overline{\theta}$. But when $\theta > \overline{\theta}$, as a member of the CU $\langle sl \rangle$, country l retains its optimal MFN tariff as a CU member (i.e., $t_l^u = t_l^\phi$) and this alters the welfare levels of the CU members and that of the non-member country.

Let $\langle \overline{sl} \rangle$ denote the CU between countries s and l when the tariff restriction of Article XXIV binds ($t_l^u = t_l^\phi$). This tariff restriction has three main implications. First, the range over which free trade is stable under Bilateralism remains unaffected. Second, the range

over which the CU between countries l and s is stable shrinks. And third, the range over which the CU between the two larger countries is stable expands.

– Figure 7 here –

As a result, the tariff constraint imposed by Article XXIV on CU members improves world welfare due to two reasons. First, when $\theta_l(F - sl) \leq \theta \leq \theta_s(\overline{sl} - \Phi)$ the CU $\langle sl \rangle$ is stable agreement with/without the tariff constraint but since the tariff protection under the constrained CU $\langle \overline{sl} \rangle$ is relatively lower, higher world welfare obtains when CU members have to abide by the tariff constraint of Article XXIV. Second, when $\theta_s(sl - \Phi) > \theta \geq \theta_s(\overline{sl} - \Phi)$, the CU $\langle sl \rangle$ is stable without the tariff constraint whereas the CU $\langle ll' \rangle$ arises when it is in place, again an outcome that is preferable from an aggregate welfare perspective.

Next, we briefly discuss how the tariff restriction of GATT Article XXIV alters the analysis under Assumption 2b. Under this assumption, the only agreement for which the tariff constraint of Article XXIV matters is the CU between the two smaller countries. Specifically, under the CU $\langle ss' \rangle$, both members are unable to impose their optimal CU tariffs ($t_s^u = t_{s'}^u$) when $\theta > \overline{\theta}$ since over this range $t_s^\phi < t_s^u$. Under such a case, each CU member imposes $t_s^u = t_{s'}^u = t_s^\phi = t_{s'}^\phi$ on the non-member. Denote this *constrained* CU by $\langle \overline{ss'} \rangle$. It is easy to show that, as under Assumption 2b, the region over which free trade is stable remains unaltered, as does the region over which the two smaller countries form a CU. Indeed, the only effect of the tariff restriction under Assumption 2b is to replace the unconstrained CU $\langle ss' \rangle$ by the constrained CU $\langle \overline{ss'} \rangle$, once again an outcome that is preferable from an aggregate welfare perspective.

– Figure 8 here –

9 Conclusion

Ever since Jacob Viner (1950) first drew attention to the issue, the economics of preferential trade agreements such as Customs Unions (CUs) has received intense scrutiny from economists and policy-makers alike. In this paper, we conduct an analysis that sheds light on whether and how the freedom to pursue CUs provided under GATT Article XXIV affects the process of multilateral trade liberalization. Our model is novel in that both the type of trade liberalization (preferential versus multilateral) and the magnitude of tariff cuts undertaken by countries under each type of liberalization are endogenously determined.

We provide an analytical comparison of two alternative games of trade liberalization. Under the Bilateralism game, countries are free to liberalize trade bilaterally via the formation of a CU, multilaterally, or not at all whereas under the Multilateralism game they can either liberalize multilaterally or not at all. This is in sharp contrast to the much of the existing literature that has tended to take either preferential trade agreements or multilateral liberalization as exogenously given. Comparing equilibrium outcomes of these two games helps provide an evaluation of GATT Article XXIV – perhaps the most important exception to the principle of non-discrimination available to WTO members today.

Our analysis shows that if governments base decisions on national welfare, no two countries would choose to form a preferential CU if endowment levels are relatively symmetric across countries. Instead, under such a scenario, global free trade would obtain as the unique stable equilibrium and the freedom to pursue discriminatory CUs would go unused. This result implies that some type of heterogeneity across countries maybe necessary to induce countries to opt out of multilateral liberalization. To verify this intuition, we consider two scenarios: one where one country is larger than the other two and another where it is smaller. Under the former scenario, the key result is that the freedom to form CUs can actually further the cause of global free trade by changing the outside

option facing the country that is reluctant to liberalize multilaterally – i.e., the smaller country. However, under the latter scenario, the relationship between the preferential and multilateral route is much more strained: the concern that CUs can serve as exclusionary devices that undermine the prospects of multilateral free trade is indeed borne out in this scenario. Finally, we also examine the impact of a key condition of Article XXIV – i.e., countries forming a CU should not raise tariffs on outsiders. Such a tariff restriction yields a more desirable CU from a global welfare perspective since it softens the adverse impact on outsiders. However, it does nothing to enhance the prospects of global free trade.

10 Appendix

In this Appendix we provide the necessary supporting calculations and proofs.

10.1 Welfare levels

We begin by reporting welfare levels under different policy regimes as a function of tariffs. Let $\mathbf{t} = (t_i, t_j, t_k)$. Since countries i and j impose MFN consistent tariffs under both the status quo and the agreement $\langle ij^m \rangle$, for a *given* vector of tariffs $\mathbf{t}$, we can write

$$w_i(\Phi) = w_i(ij^m) = w_i(jk^m) = CS_i(ij^m) + PS_i(ij^m) + TR_i(ij^m) \tag{44}$$

where

$$CS_i(ij^m) = \frac{1}{18}\left[(e_i + e_j + t_k)^2 + (e_i + e_k + t_j)^2 + (e_j + e_k - 2t_i)^2\right]$$

and

$$PS_i(ij^m) = e_i \left(2\alpha - \frac{2e_i + \sum_{z \neq i}(e_z + t_z)}{3} \right)$$

and

$$TR_i(ij^m) = \frac{t_i}{3}\left(\sum_{z\neq i} e_z - 2t_i\right)$$

Under the status quo $\langle\Phi\rangle$ country i chooses its non-discriminatory tariff t_i to maximize its own welfare. We have:

$$\frac{\partial w_i(\Phi)}{\partial t_i} = \frac{e_j + e_k - 8t_i}{9} = 0 \Longrightarrow t_i^{\phi} = \frac{e_j + e_k}{8} \tag{45}$$

Note further that $t_i\left(jk^m\right) = t_i^{\phi}$ since the excluded country under the MFN-consistent agreement $\langle jk\rangle^m$ continues to impose its optimal MFN tariff on its trading partners.

Under the agreement $\langle ij\rangle^m$, countries i and j jointly choose their non-discriminatory tariffs $(t_i^m,\ t_j^m)$ to maximize $w_i(ij^m) + w_j(ij^m)$. The relevant first order conditions and the associated optimal tariffs are:

$$\begin{aligned}
\frac{\partial(w_i(ij^m) + w_j(ij^m))}{\partial t_i} &= \frac{2e_k - e_j - 7t_i}{9} = 0 \Longrightarrow t_i^m = \frac{2e_k - e_j}{7} \\
\frac{\partial(w_i(ij^m) + w_j(ij^m))}{\partial t_j} &= \frac{2e_k - e_i - 7t_j}{9} = 0 \Longrightarrow t_j^m = \frac{2e_k - e_i}{7}
\end{aligned} \tag{46}$$

Under the CU $\langle ij\rangle$, members abolish tariffs on each other $(t_{ij} = t_{ji} = 0)$ and impose jointly optimal tariffs on the non-member. Let t_i and t_j denote member countries' external tariffs and t_k denote the non-member's non-discriminatory tariff. Then, the welfare of member i equals

$$w_i(ij) = CS_i(ij) + PS_i(ij) + TR_i(ij) \tag{47}$$

where

$$CS_i(ij) = \frac{1}{18}\left[(e_i + e_j + t_k)^2 + (e_i + e_k - t_j)^2 + (e_j + e_k - t_i)^2\right]$$

and

$$PS_i(ij) = e_i \left(2\alpha - \frac{2e_i + \sum_{z \neq i} e_z + t_k - t_j}{3} \right)$$

and

$$TR_i(ij) = \frac{t_i}{3} (2e_k - e_j - 2t_i)$$

Member countries choose their external tariffs to maximize their joint welfare $w_i(ij) + w_j(ij)$. This gives:

$$\begin{aligned} \frac{\partial (w_i(ij) + w_j(ij))}{\partial t_i} &= \frac{4e_k - 2e_j - 10t_i}{9} = 0 \Longrightarrow t_i^u = \frac{2e_k - e_j}{5} \\ \frac{\partial \left(w_i(ij) + w_j(ij) \right)}{\partial t_j} &= \frac{4e_k - 2e_i - 10t_j}{9} = 0 \Longrightarrow t_j^u = \frac{2e_k - e_i}{5} \end{aligned} \tag{48}$$

Now, consider the welfare of a non-member country i under $\langle jk \rangle$:

$$w_i(jk) = CS_i(jk) + PS_i(jk) + TR_i(jk) = w_i(\Phi)$$

where $CS_i(jk)$. Thus, the non-member country i's tariff under the CU $\langle jk \rangle$ equals t_i^ϕ:

$$\frac{\partial w_i(jk)}{\partial t_i} = \frac{e_j + e_k - 8t_i}{9} = 0 \Longrightarrow t_i\left(jk\right) = t_i^\phi = \frac{e_j + e_k}{8}$$

Finally, by imposing $t_i = t_j = t_k = 0$ in (44), we obtain $w_i^*(F)$ – the welfare of country i under free trade:

$$w_i^*(F) = w_i(\Phi)|_{t_i = t_j = t_k = 0} \tag{49}$$

Using the welfare and tariff levels reported above, we can easily obtain the formulae for welfare levels under all possible regimes. Lemmas 2 and 3 and the various inequalities reported in the main text follow from a direct application of the relevant formulae. For example, consider the world welfare comparisons reported in (39) and (40). A direct

application of the relevant formulae yields: $\Delta ww^*(ll'^m - sl) \geq 0$ iff $\theta \leq \frac{5}{3}$. Indeed, the same parameter restriction (i.e., $\theta \leq \frac{5}{3}$) ensures that (1) $\Delta ww^*(ll'^m - ll') \geq 0$; (2) $\Delta w_s^*(ll'^m - sl) \geq 0$; (3) $\Delta w_{l'}^*(ll'^m - sl) \geq 0$; (4) $\Delta ww^*(sl^m - ss') \geq 0$. To save space, we do not include the algebraic details underlying these straightforward calculations.

10.2 Inequality (26)

We have $x_j^I(\Phi) = x_j^I(jk) = x_j^I(jk^m) = \frac{5e_j - 3e_k}{8} \geq 0$ iff $e_k \leq \frac{5e_j}{3}$. Furthermore, under the CU $\langle ij \rangle$ we have $x_j^I(ij) = \frac{3e_j - e_k}{5} \geq 0$ iff $e_k \leq 3e_j$ and $x_j^I(ik) = \frac{2e_j - e_k}{5} \geq 0$ iff $e_k \leq 2e_j$. Next, under the MFN consistent agreement $\langle ij \rangle^m$ we have, $x_j^I(ij^m) = \frac{5e_j - 3e_k}{7} \geq 0$ iff $e_k \leq \frac{5e_j}{3}$ and $x_j^I(ik^m) = \frac{4e_j - 2e_k}{7} \geq 0$ iff $e_k \leq 2e_j$. Finally, under free trade, we have $x_j^I(F) = \frac{2e_j - e_k}{3} \geq 0$ iff $e_k \leq 2e_j$.

10.3 Proof of Proposition 3

Consider first the CU $\langle ll' \rangle$. There are three possible defections from it:

(ULL) Unilateral deviation of l or l' from $\langle ll' \rangle$ to $\langle \Phi \rangle$.

(JLL1) Joint defection of s and l (or $l\prime$) from $\langle ll' \rangle$ to $\langle sl \rangle$ (or $\langle sl' \rangle$)

(JLL2) Joint defection of s, l and l' from $\langle ll' \rangle$ to $\langle F \rangle$

We know from (14) that under symmetry no member country has an incentive to unilaterally deviate from a CU to the status quo. From Lemma 2, we have $\frac{\partial \Delta w_l^*(ll' - \Phi)}{\partial \theta} = \frac{\partial \Delta w_{l'}^*(ll' - \Phi)}{\partial \theta} < 0$. Next note that

$$\Delta w_l^*(ll' - \Phi) = \Delta w_{l'}^*(ll' - \Phi) > 0 \text{ when } \theta = \frac{5}{3}$$

Thus, we can argue that (ULL) never occurs:

$$\Delta w_l^*(ll' - \Phi) = \Delta w_{l'}^*(ll' - \Phi) > 0 \text{ for all } \theta \tag{50}$$

Now consider (JLL1). From (27), we know that a larger country (say l) always has an

incentive to deviate from $\langle ll' \rangle$ to $\langle sl \rangle$ while the smaller country deviates only when the degree of endowment asymmetry is not large:

$$\Delta w_s^*(sl - ll') \geq 0 \text{ iff } \theta \leq \theta_s(sl - ll') \cong 1.28 \tag{51}$$

and that this joint deviation is self-enforcing when $\theta \leq \theta_s(sl - \Phi) \cong 1.27$.

Now consider (JLL2). It is immediate from (32) and (33) that (JLL2) occurs when $\theta \leq \theta_s(F - ll') \cong 1.23$ and it is self enforcing iff $\theta \leq \theta_s(F - sl) \cong 1.048$. These deviations together imply that $\langle ll' \rangle$ is stable when $\theta \geq \theta_s(sl - \Phi)$.

Now consider the stability of $\langle sl \rangle$. Note from (36) that country s has an incentive to unilaterally deviate from $\langle sl \rangle$ to $\langle \Phi \rangle$ when $\theta > \theta_s(sl - \Phi)$. It is straightforward from (27) that there exist no joint deviations from $\langle sl \rangle$ to $\langle ll' \rangle$. Furthermore, all countries deviate to free trade when $\theta < \theta_l(F - sl) \cong 1.045$ and this joint deviation is self-enforcing. As a result, $\langle sl \rangle$ or $\langle sl' \rangle$ is stable when $\theta_l(F - sl) \leq \theta \leq \theta_s(sl - \Phi)$.

Finally, we examine the stability of the status quo under Bilateralism. Note from (50) that two larger countries jointly deviate from $\langle \Phi \rangle$ to $\langle ll' \rangle$ and this deviation is self-enforcing. Thus $\langle \Phi \rangle$ is not stable.

10.4 Critical θ values in Propositions 4, 5, 6, 8 and 9

In the multilateralism game, $\theta_s(F - ll'^m) \cong 1.015$ obtains under Assumption 2a while $\theta_{s'}(F - sl^m) \cong 1.03$ and $\theta_s(sl^m - \Phi) \cong 1.15$ obtain under Assumption 2b.

10.5 Proof of Lemma 4

It is immediate from (16) and Lemma 2 that

$$\Delta w_l^*(F - ss') > 0 \text{ for all } \theta \tag{52}$$

Moreover, we know from Lemma 2 that $\frac{\partial \Delta w_l^*(F-sl)}{\partial \theta} > 0$. Then, combining this with inequality (17) yields

$$\Delta w_l^*(F - sl) > 0 \text{ for all } \theta \tag{53}$$

Finally, combining (14) and Lemma 2 implies that $\Delta w_l^*(sl - \Phi) > 0$. Then, from this inequality and (53) it is immediate that $\Delta w_l^*(F - \Phi) > 0$.

10.6 Proof of Proposition 7

In light of Lemma 4, we focus on the perspective of countries s and s'. Direct calculations yield

$$\Delta w_s^*(F - ss') \leq 0 \text{ iff } \theta \geq \theta_s(F - ss') \cong 1.02 \tag{54}$$

Moreover, Lemma 2 and (14) together implies that

$$\Delta w_s^*(ss' - \Phi) > 0 \text{ for all } \theta \tag{55}$$

Thus, the joint deviation from $\langle F \rangle$ to $\langle ss' \rangle$ is self-enforcing. Inequality (55) also implies that the status quo is not stable since the smaller countries jointly deviate from $\langle \Phi \rangle$ to $\langle ss' \rangle$ and this deviation is self-enforcing.

Further note that

$$\Delta w_s^*(F - \Phi) \leq 0 \text{ iff } \theta \geq \theta_s(F - \Phi) \cong 1.39 \tag{56}$$

However, $\theta_s(F - \Phi) > \theta_s(F - ss')$ and thus the above deviation is not binding for the stability of $\langle F \rangle$. Finally, note from (14), (15) and Lemma 2 that the smaller countries have an incentive to jointly deviate from $\langle sl \rangle$ to $\langle ss' \rangle$ and this deviation is a self-enforcing deviation – see (55). Thus, the CU between a small country and the large country $\langle sl \rangle$ is not stable.

10.7 When the tariff restriction of Article XXIV binds

Under assumption 2a

First, consider the stability of free trade. To this end, note that when $\theta > \overline{\theta}$, we have

$$\Delta w_l^*(F - \overline{sl}) < 0 \text{ and } \Delta w_s^*(F - \overline{sl}) < 0 \text{ when } \theta > \overline{\theta} \tag{57}$$

and this joint deviation is self-enforcing only when $\theta \leq \theta_s(\overline{sl} - \Phi)$; otherwise country s has an incentive to further deviate unilaterally from $\langle \overline{sl} \rangle$ to $\langle \Phi \rangle$:

$$\Delta w_s^*(\overline{sl} - \Phi) \leq 0 \text{ iff } \theta \geq \theta_s(\overline{sl} - \Phi) = \frac{5}{4} \tag{58}$$

Combining this with (30), (32) and (33), we find that the region over which free trade is stable does not change relative to the case where the CU between countries s and l is free to raise its tariffs on the non-member.

Next we examine the stability of $\langle \overline{sl} \rangle$. As before, we have

$$\Delta w_l^*(ll' - \overline{sl}) < 0 \tag{59}$$

Thus, it is immediate from (57), (58) and (59) that $\langle \overline{sl} \rangle$ is stable when $\theta_l(F - sl) \leq \theta \leq \theta_s(\overline{sl} - \Phi)$. Note that since $\theta_s(\overline{sl} - \Phi) < \theta_s(sl - \Phi)$, the region over which CU between countries s and l is stable is smaller when Article XXIV binds ($t_l^u = t_l^\phi$). Finally note that countries s and l have no joint incentive to deviate from $\langle ll' \rangle$ to $\langle \overline{sl} \rangle$ when θ is sufficiently small:

$$\Delta w_s^*(ll' - \overline{sl}) \leq 0 \text{ iff } \theta \leq \theta_s(\overline{sl} - ll') \cong 1.26 \tag{60}$$

and this deviation is self-enforcing only when $\theta \leq \theta_s(\overline{sl} - \Phi)$. If we combine this result with (32), (33) and (50), it follows that $\langle ll' \rangle$ is stable when $\theta \geq \theta_s(\overline{sl} - \Phi)$. Thus, the region over which the CU between countries l and l' is stable expands if members cannot raise tariffs on the outsider.

Under assumption 2b

First, consider the stability of free trade. To this end, note that when $\theta > \overline{\theta}$, the following holds:

$$\Delta w_s^*(F - \overline{ss'}) = \Delta w_{s'}^*(F - \overline{ss'}) < 0 \tag{61}$$

Furthermore, we have

$$\Delta w_s^*(\overline{ss'} - \Phi) = \Delta w_{s'}^*(\overline{ss'} - \Phi) > 0 \tag{62}$$

Thus, it is immediate that the region over which free trade is stable does not change with the binding tariff constraint. Moreover, it is clear from (55) and (62) that the status quo is not stable. Finally note that, the joint deviation from $\langle \overline{ss'} \rangle$ to $\langle sl \rangle$ does not occur. This analysis implies that $\langle \overline{ss'} \rangle$ is stable when $\theta \geq \theta_s(F - ss')$ and therefore the region over which CU between two smaller countries is stable remain unaffected by the tariff constraint imposed by Article XXIV.

However, when $\theta \geq \theta_s(F - ss')$, the CU $\langle ss' \rangle$ arises as the stable agreement without the constraint while $\langle \overline{ss'} \rangle$ obtains with the constraint. Since the total tariff protection is lower (thus trade diversion is smaller) under the constrained CU $\langle \overline{ss'} \rangle$ relative to the unconstrained $\langle ss' \rangle$, higher world welfare obtains under the constrained CU.

References

[1] Aghion, Philippe, Pol Antràs, and Elhanan Helpman, 2007. "Negotiating Free Trade." Journal of International Economics 73, 1-30.

[2] Bagwell, Kyle and Robert. W. Staiger, 1997a. "Multilateral Tariff Cooperation During the Formation of Customs Unions." Journal of International Economics, Elsevier 42, 91-123.

[3] Bagwell, Kyle and Robert. W. Staiger, 1997b. "Multilateral Tariff Cooperation During the Formation of Free Trade Areas." International Economic Review 38, 291-319.

[4] Bagwell, Kyle and Robert. W. Staiger, 1998. "Regionalism and Multilateral Tariff Cooperation." In John Piggott and Allan Woodland, eds, International Trade Policy and the Pacific Rim, London: MacMillan.

[5] Bernheim, Douglas B., Bezalel Peleg and Michael Whinston, 1987. "Coalition-proof Nash Equilibria I. Concepts." Journal of Economic Theory 42, 1-12.

[6] Bhagwati, Jagdish. The World Trading System at Risk, 1991, Princeton University Press, Princeton, NJ.

[7] Bhagwati, Jagdish, Arvind Panagariya, and Pravin Krishna, eds., Trading Blocs, 1999, The MIT Press, Cambridge, MA.

[8] 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.

[9] 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.

[10] Estevadeordal, Antoni, Caroline Freund and Emanuel Ornelas, 2008. " Does Regionalism Affect Trade Liberalization toward NonMembers" The Quarterly Journal of Economics, 123(4), 1531-1575.

[11] Freund, Caroline, 2000. "Multilateralism and the Endogenous Formation of Preferential Trade Agreements." Journal of International Economics 52, 359-376.

[12] Furusawa, Taiji, 2002. "Preferential Trade Arrangements vs. Open. Regionalism: A Theoretical Analysis of APEC" In Robert M. Stern, eds, Issues and Options for US-Japan Trade Policies, The University of Michigan Press.

[13] Furusawa, Taiji and Hideo Konishi, 2007. "Free Trade Networks." Journal of International Economics 72, 310-335.

[14] Gatsios, Konstantine and Karp, Larry, 1991. "Delegation in a Customs Union," The Review of Economic Studies 58, 391-397

[15] Goyal, Sanjeev and Sumit Joshi, 2006. "Bilateralism and Free Trade." International Economic Review 47, 749-778.

[16] Grossman, Gene M. and Elhanan Helpman, 1995. "The Politics of Free-Trade Agreements." American Economic Review 85, 667-690.

[17] Hadjiyiannis, Costas, 2004. "Regionalism and Multilateral Trade Liberalization with Asymmetric Countries," Review of International Economics, 12, 395-411.

[18] Kennan, John and Raymond Riezman, 1990. "Optimal Tariff Equilibria with Customs Unions?" Canadian Journal of Economics 23, 70-83.

[19] Krishna, Pravin, 1998. "Regionalism and Multilateralism: A Political Economy Approach." The Quarterly Journal of Economics 113, 227-251.

[20] Melatos, Mark and Alan Woodland, 2007. "Endogenous Trade Bloc Formation in an Asymmetric World" European Economic Review 51, 901-924.

[21] Mrazova, Monika, David Vines and Ben Zissimos, 2010. "Is the WTO's Article XXIV Bad?" unpublished manuscript.

[22] Ornelas, Emanuel, 2005a. "Trade Creating Free Trade Areas and the Undermining of Multilateralism." European Economic Review 49, 1717-1735.

[23] Ornelas, Emanuel, 2005b. "Rent Destruction and the Political Viability of Free Trade Agreements." Quarterly Journal of Economics 120, 1475-1506.

[24] Raimondos-Møller, Pascalis and Alan D. Woodland, 2006. "Non-preferential Trading Clubs." Journal of International Economics 68, 79-91.

[25] Riezman, Raymond, 1985. "Customs Unions and the Core?" Journal of International Economics 19, 355-365.

[26] Riezman, Raymond, 1999. "Can Bilateral Trade Agreements Help Induce Free Trade?" Canadian Journal of Economics 32, 751-766.

[27] Saggi, Kamal, 2006. "Preferential Trade Agreements and Multilateral Tariff Cooperation." International Economic Review 47, 29-57.

[28] Saggi, Kamal and Halis M. Yildiz, 2010. "Bilteralism, Multilateralism, and the Quest for Global Free Trade." Journal of International Economics 81, 26-37.

[29] Seidmann, Daniel, 2009. "Preferential Trading Arrangements as Strategic Positioning" Journal of International Economics 79, 143-159.

[30] Syropoulos, Constantinos, 2003. "Rules for the Disposition of Tariff Revenues and the Determination of Common External Tariffs in Customs Unions." Journal of International Economics 60, 387-416.

[31] Viner, Jacob, 1950. The Custom Union Issue. New York: Carnegie Endowment for International Peace.

[32] Yi, Sang-Seung, 1996. "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good." Journal of International Economics 41, 153-177.

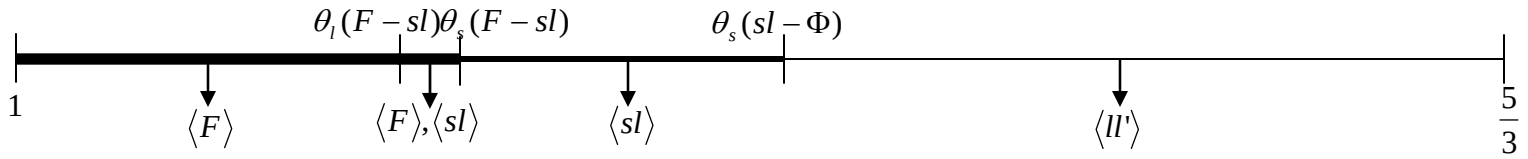

Figure 1: Stable agreements under Bilateralism: one country is smaller

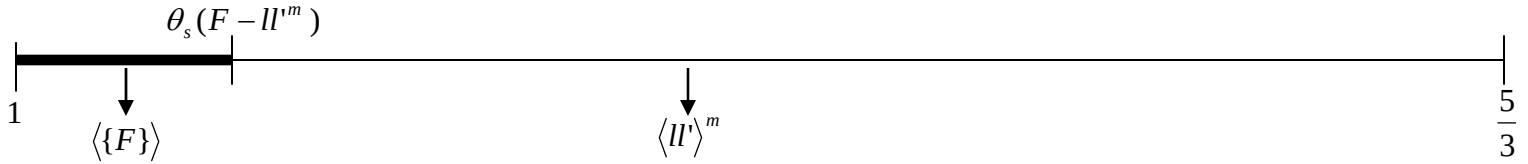

Figure 2: Stable agreements under Multilateralism: one country is smaller

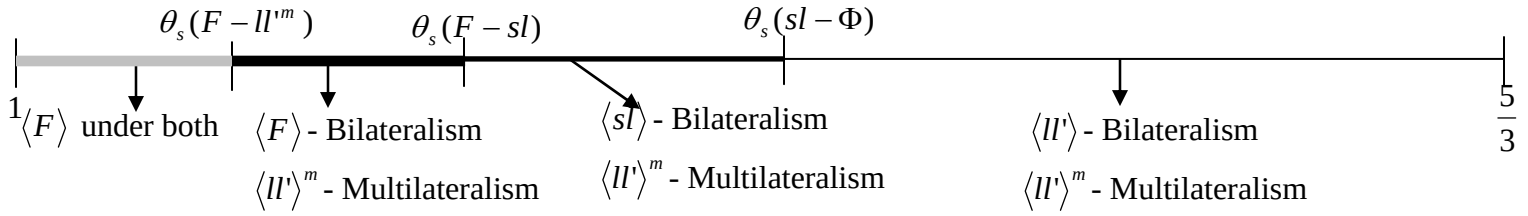

Figure 3: Bilateralism versus Multilateralism: one country is smaller

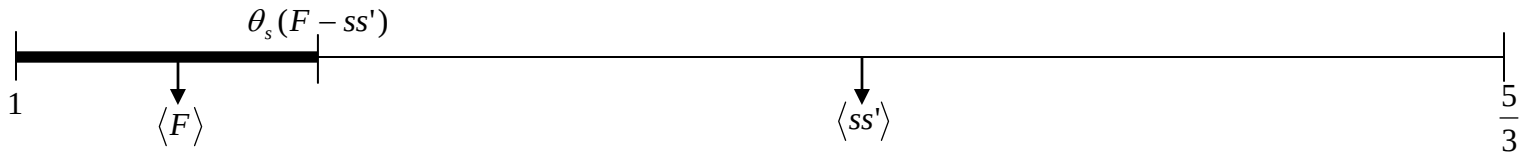

Figure 4: Stable agreements under Bilateralism: one country is larger

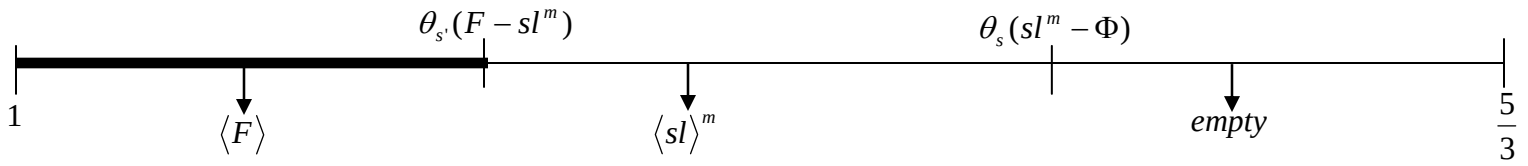

Figure 5: Stable agreements under Multilateralism: one country is larger

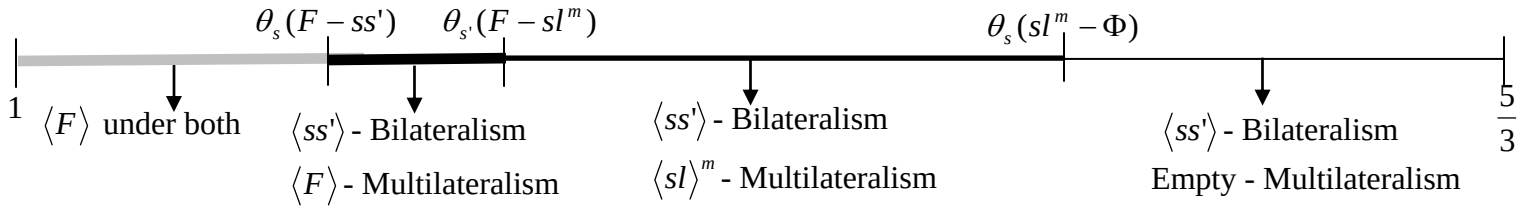

Figure 6: Bilateralism versus Multilateralism: one country is larger

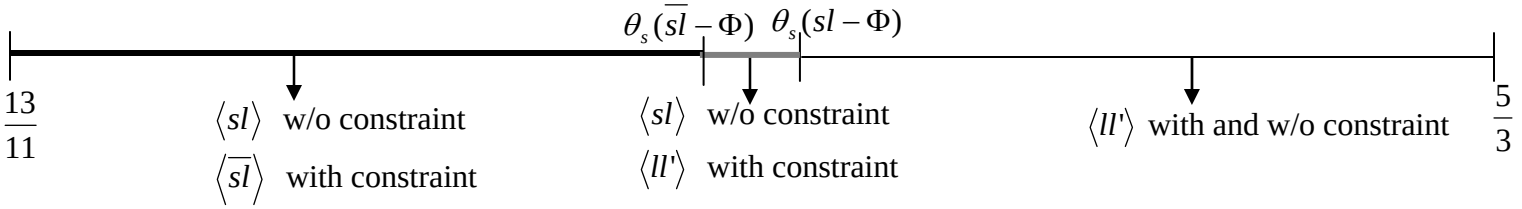

Figure 7: Binding tariff constraint of Article XXIV: one country is smaller

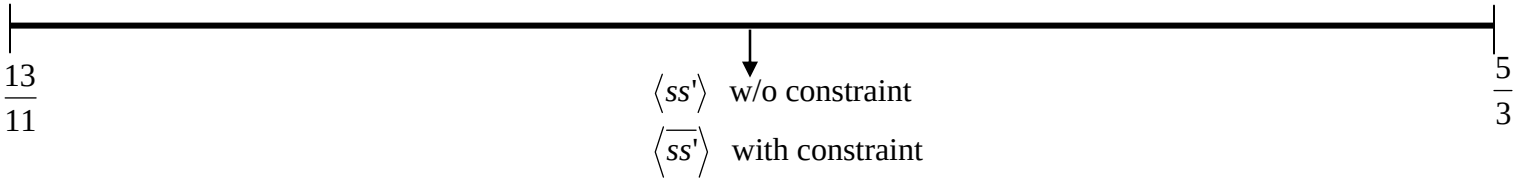

Figure 8: Binding tariff constraint of Article XXIV: one country is larger