Bilateralism, multilateralism, and the quest for global free trade

Saggi, Kamal and Yildiz, Halis Murat
Southern Methodist University, Ryerson University

30. June 2009

Bilateralism, multilateralism, and the quest for global free trade*

Kamal Saggi† and Halis Murat Yildiz‡

Abstract

We develop an equilibrium theory of trade agreements in which both the *degree* and the *nature* (bilateral or multilateral) of trade liberalization are endogenously determined. To determine whether and how bilateralism matters, we also analyze a scenario where countries pursue trade liberalization on only a multilateral basis. We find that when countries have asymmetric endowments or when governments value producer interests more than tariff revenue and consumer surplus, there exist circumstances where global free trade is a stable equilibrium only if countries are free to pursue bilateral trade agreements. By contrast, under symmetry, both bilateralism and multilateralism yield global free trade.

Keywords: Bilateral trade agreements, multilateral trade liberalization, free trade agreements, GATT. JEL Classifications: F13, F12.

---

*For helpful comments, we thank seminar audiences at the Asia-Pacific Seminar at the University of Sydney, Fourth Annual Development Conference of the Indian Statistical Institute, George Washington University, Graduate School of Economics at Hitotsubashi University, IEFS-China Inaugural Conference at the University of International Business and Economics, Louisiana State University, National University of Singapore, SMU, University of Guelph, University of Missouri, World Bank, and the WTO. We also thank Kyle Bagwell, Rick Bond, Theresa Carpenter, Caroline Freund, Bernard Hoekman, Giovanni Maggi, Aaditya Mattoo, Andres Rodriguez-Clare, Bob Staiger, Alan Woodland, and two anonymous referees for many constructive comments and suggestions. Any remaining errors are our own.

†Department of Economics, Southern Methodist University, Dallas, TX 75275-0496. Phone: 214-768-3274; fax: 214-768-1821; e-mail: ksaggi@smu.edu. The author thanks the World Bank for financial support.

‡Department of Economics, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. Phone: 416-979-5000 (ext 6689); fax: 416-979-5289; e-mail: hyildiz@ryerson.ca. The author gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC).

1 Introduction

Global trade liberalization occurs through a variety of channels, not all of which appear to be in harmony with one another. While practically every major nation is now a member of the World Trade Organization (WTO) and a participant in its complex process of multilateral trade liberalization, an average WTO member also belongs to six preferential trade agreements (PTAs) (World Bank, 2005). The schizophrenic nature of today's multilateral trading system is reflected in the somewhat conflicting rules of the WTO's key multilateral trade agreement, i.e. the General Agreement for Tariffs and Trade (GATT): while Article I of GATT requires member countries to undertake trade liberalization on a most-favored-nation (MFN) or non-discriminatory basis, Article XXIV of the very same agreement permits WTO member countries to pursue PTAs under which participating countries grant tariff (and other trade policy) concessions to each other that they do *not* have to extend to all member countries of the WTO.[1] This raises the following question: would GATT serve the cause of global free trade more effectively if it did *not* include the exception to MFN provided by Article XXIV? In other words, would global free trade be easier to achieve if all WTO members were to pursue trade liberalization on *only* a multilateral basis? To address this issue, we develop an equilibrium theory of trade agreements and use it to compare the pros and cons of bilateral and multilateral approaches to trade liberalization.

We analyze the coalition proof (or stable) Nash equilibria of a game of trade liberalization between three countries that differ with respect to their endowment levels. The game (which we refer to as *bilateralism*) proceeds as follows. In the first stage, each country announces the names of its trading partner(s) with whom it wishes to sign a free trade agreement (FTA). An FTA between two countries requires them to abolish tariffs on each other and it arises iff they both announce each other's name. Next, given the world trade regime, countries choose their tariffs. Finally, international trade and consumption take place. After analyzing equilibrium trade agreements under bilateralism, we examine the stable equilibria of this game under the

[1] While Article XXIV tries to limit the damage on non-member countries by requiring PTA members to not raise tariffs on outsiders, the fact remains that it contradicts the principle of non-discrimination that underlies the entire WTO system.

restriction that countries can liberalize trade on *only* a multilateral basis (we call this game *multilateralism*). By comparing equilibrium outcomes under bilateralism with those under multilateralism, we isolate the consequences of the exception to multilateral trade liberalization available to WTO members under GATT Article XXIV.[2] To the best of our knowledge, our paper is the first to provide such a comparison in a model in which both the *nature* and the *degree* of trade liberalization are endogenously determined.

Consistent with actual WTO experience, under our multilateralism game a pair of countries can engage in mutual trade liberalization so long as each of them extends their respective tariff reductions also to the third country – as mandated by the GATT's MFN clause.[3] We find that the degree of trade liberalization undertaken by two countries (say i and j) under such a multilateral trade agreement $\langle\{ij^m\}\rangle$ is lower relative to that under the bilateral free trade agreement $\langle\{ij\}\rangle$. As a result, the non-participating country (i.e. k) actually faces *lower* tariffs in export markets under the bilateral FTA $\langle\{ij\}\rangle$ relative to the multilateral trade agreement $\langle\{ij^m\}\rangle$. However, the non-member country faces discriminatory tariffs in export markets under $\langle\{ij\}\rangle$ whereas no such discrimination exists under $\langle\{ij^m\}\rangle$. Due to this crucial difference between the two types of trade agreements, in our model the non-member country is worse off under the bilateral FTA $\langle\{ij\}\rangle$ relative to the multilateral agreement $\langle\{ij^m\}\rangle$.

Equilibrium analysis reveals that when countries are symmetric with respect to their endowment levels, global free trade is the only stable equilibrium under both bilateralism and multilateralism. In other words, under symmetry, the freedom to pursue purely bilateral agreements has no consequences at all insofar as the obtainment of global free trade is concerned. This immediately raises the question whether this irrelevance result holds when the underlying economic environment is asymmetric in some respects. To this end, we then analyze a scenario where endowment levels are unequal across countries and find that global free trade is stable over a larger parameter space under bilateralism relative to multilateralism. This result has a powerful and surprising implication: there exist circumstances where global free trade is a

[2]We do not consider unilateral trade liberalization since the presence of terms of trade considerations in our model implies that such liberalization is not in any country's interest.

[3]Note that in our model countries are free to sign a multilateral agreement even under bilateralism. By contrast, the multilateralism game rules out a discriminatory bilateral FTA.

stable equilibrium *only if* countries are free to form bilateral FTAs. Why? The logic is as follows. While considering whether or not to participate in multilateral trade liberalization, each country has to take into account its welfare under the trade regime that emerges in the *absence* of its participation. Since a non-participating country (say k) is worse off under the bilateral FTA $\langle\{ij\}\rangle$ relative to the multilateral agreement $\langle\{ij^m\}\rangle$, each country's incentive to participate in multilateral trade liberalization is stronger when its non-participation results in a discriminatory bilateral FTA between the other two countries as opposed to when it results in a non-discriminatory multilateral agreement between them.[4] In this way, the freedom to pursue bilateral agreements can act as a force in favor of multilateral trade liberalization.[5]

Our results show that heterogeneity across countries may be an important determinant of the potential for success of multilateralism and that bilateralism has a useful role to play in the process of global trade liberalization. An important implication of our analysis is that to properly account for the role of bilateralism, we need to better understand *why* countries choose to enter into bilateral agreements when multilateral trade liberalization is an option. In this context it is noteworthy that while both Krugman (1991) and Grossman and Helpman (1995) noted that asymmetries across countries can play a crucial role in determining incentives for bilateral and multilateral trade liberalization, existing literature has tended to pay little attention to this issue. Indeed, in our model bilateral FTAs even fail to arise when countries have symmetric endowments since, under such circumstances, countries find it in their mutual interest to go all the way to global free trade.

In a recent paper, Aghion et. al. (2007) examine a leading country's choice between sequential and multilateral bargaining of free trade agreements and provide a comparison of these bargaining processes. While we consider similar issues, there are important differences

[4]Since political economy considerations can potentially play an important role in determining incentives for bilateral and multilateral trade liberalization, later in the paper (section 6) we consider a scenario where governments put greater weight on producer interests relative to consumer surplus and tariff revenue. We find that the presence of such political economy motives actually enlarges the parameter space over which the freedom to pursue bilateral trade agreements is necessary for achieving global free trade.

[5]Saggi and Yildiz (2006) consider cost differences across countries in an oligopolistic model of intraindustry trade and uncover similar results. See Levy (1997), Krishna (1998), and Ornelas (2005b) for analyses focusing on political economy considerations.

between their approach and ours. First, in our model, *all* countries are free to negotiate FTAs and countries can form multiple bilateral FTAs. Second, under multilateralism, Aghion et al. (2007) assume that countries have only a binary choice between global free trade and no agreement whereas we permit two countries to undertake trade liberalization on an MFN basis. In Aghion et al. (2007), when bilateral FTAs are forbidden (i.e. under their multilateral bargaining protocol), any single country (say country k) can ensure that no agreement $\langle\{\Phi\}\rangle$ prevails by simply opting to not practise free trade itself while such a country ends up facing the multilateral agreement $\langle\{ij^m\}\rangle$ in our model. As a result, while the nature of coalition externalities (i.e. whether they are negative or positive) shapes the choice between sequential and multilateral bargaining and the circumstances under which global free trade obtains in their model, the *relative degree of the positive externality* under the bilateral FTA $\langle\{ij\}\rangle$ and the multilateral agreement $\langle\{ij^m\}\rangle$ is the driving force behind our analysis. More specifically, while both $\langle\{ij\}\rangle$ and $\langle\{ij^m\}\rangle$ generate a positive externality for the non-member by increasing its welfare relative to the status quo, the degree of the positive externality is stronger under the multilateral agreement $\langle\{ij^m\}\rangle$. Finally, unlike Aghion et. al. (2007), we do not allow transfers between different coalitions. This is important because when transfers are possible and global free trade maximizes aggregate welfare, it necessarily emerges as the equilibrium under both sequential and multilateral bargaining. When free trade does not maximize aggregate welfare, Aghion et. al. (2007) find that FTAs facilitate the achievement of global free trade iff they create negative externalities for non-members. In our model, FTAs can have this effect even when free trade maximizes global welfare and FTAs generate a positive externality for the non-member.

Our paper shares some key elements with Goyal and Joshi (2006) and Furusawa and Konishi (2007), both of which employ the network formation game developed by Jackson and Wolinsky (1996) in examining whether or not a given trade configuration is pair-wise stable.[6] Under symmetry, global free trade is also stable under their approach. Unlike, us however,

[6]Relative to our approach, the concept of pairwise stability implies two constraints. First, the deviating coalition can contain at most two countries. Second, the deviation can consist of severing just one existing link or forming one additional link. In order to eliminate these constraints, we follow Bernheim et al. (1987) and use the concept of coalition proof Nash equilibrium to isolate stable equilibria.

they only examine whether the formation of bilateral FTAs results in global free trade as the stable outcome and do not analyze the consequences of adopting a strictly multilateral approach to trade liberalization. The approach of this paper is also related to that of Riezman (1999) who also asks whether bilateralism facilitates or hinders the achievement of global free trade. However, while we analytically derive the stable Nash equilibria of a non-cooperative game of FTA formation, Riezman (1999) uses the cooperative solution concept of the core and illustrates his results via numerical examples. Second, our model allows us to focus on asymmetries across countries in a way that cannot be done in Riezman's (1999) framework. The relationship between preferential and multilateral liberalization has frequently been analyzed in the literature in models of repeated interaction between countries – see Bagwell and Staiger (1997), Bond et. al. (2001), Freund (2000), and Saggi (2006).[7] We add value to this literature by treating both bilateral and multilateral liberalization as endogenous.[8]

2 Underlying trade model

To endogenize the formation of trade agreements among asymmetric countries, we utilize an appropriately adapted version of the partial equilibrium framework developed by Bagwell and Staiger (1999). 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 where $e_a \leq e_b \leq e_c$.[9]

The demand for good z in country i is given by $d(p_i^z) = \alpha - p_i^z$ where $z = A, B$, or C. As is well known, the above demand functions can be derived from a utility function of the form $U(c^z) = u(c^z) + w$ where c^zdenotes consumption of good z; w denotes the numeraire good; and $u(c^z)$ is quadratic and additively separable in each of the three goods. Since each

[7] See Bhagwati et. al. (1999) for a collection of many of the important papers in the area.

[8] Note that while the joint welfare of members is often considered in determining whether or not a trade agreement would arise (see, for example, Ornelas (2005a) and Krishna (1998)), much of the existing literature on free trade agreements does not derive the equilibria of a game that fully specifies the process by which such agreements arise.

[9] In addition, all countries have large enough endowments of the numeraire good w to ensure trade balance.

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:

$$p_i^I = p_j^I + t_{ij} \ = p_k^I + t_{ik} \tag{1}$$

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

$$m_i^I = d(p_i^I) = \alpha - p_i^I \tag{2}$$

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

$$x_j^I = e_j - [\alpha - p_j^I] \tag{3}$$

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

$$m_i^I = \sum_{j \neq i} x_j^I \tag{4}$$

Equations (1) through (4) imply that the equilibrium price of good I in country i equals:

$$p_i^I = \frac{1}{3}\left(3\alpha - \sum_{j \neq i} e_j + \sum_{j \neq i} t_{ij}\right) \tag{5}$$

Using these prices, the volume of trade is easily calculated. As is clear from equation (5), 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 (5): only a third of a given increase in either of its tariffs is passed on to domestic

consumers with exactly two third of the increase falling on the shoulders of foreign exporters.

By design the model examines country i's trade protection towards only good I (i.e. the only non-numeraire good that it imports). Since countries have asymmetric endowments, under free trade country a faces the largest volume of imports of protected goods (it imports $(e_b+e_c)/3$ units of good A) whereas country c faces the lowest volume of imports of such goods (it imports $(e_a+e_b)/3$ units of good C).[10] Note also that country i's imports of good I do not equal its exports of other non-numeraire goods. For example, under free trade, country a exports $(2e_a-e_b)/3$ units of good C to country c and $(2e_a-e_c)/3$ units of good B to country b and the sum of these exports is lower than its total imports of good C: $0 < 4e_a-e_b-e_c < e_b+e_c$. In order to balance trade, in addition to exporting goods B and C, country a exports the numeraire good to both countries b and c. Similarly, country c imports the numeraire good from both its trading partners.

From a welfare perspective, given the partial equilibrium nature of the model, it suffices to consider only protected goods. A country's welfare is defined as the sum of consumer surplus, producer surplus, and tariff revenue over all such goods:

$$w_i = \sum_z CS_i^z + \sum_z PS_i^z + TR_i \tag{6}$$

Using equations (1) through (5) one can easily obtain welfare of country i as a function of endowment levels and tariffs. Let aggregate world welfare be defined as the sum of each country's welfare: $ww = \sum_i w_i$.

We proceed as follows. First, we consider a three stage game of trade liberalization under which each country is free to pursue either (a) no trade liberalization or (b) bilateral trade liberalization or (c) multilateral trade liberalization.[11] This game is meant to capture the various options regarding trade liberalization available to WTO members, option (b) being made possible by GATT Article XXIV. After deriving Nash equilibria and isolating those equilibria

[10] The same ranking applies with respect to the value of imports so long as $3\alpha > e_a + e_b + 2e_c$, which is a minor condition that is assumed to hold.

[11] Since all countries have market power in our model, allowing for unilateral liberalization is not necessary: no country will choose to pursue such liberalization in our model.

that are *stable* (more on this below), we ask how equilibrium outcomes are affected if countries cannot pursue bilateral trade liberalization and must instead choose between multilateral trade liberalization or no liberalization at all. This exercise allows to isolate the consequences of the exception to MFN provided under GATT Article XXIV.

3 Endogenous trade agreements

We now describe our game of trade liberalization (which we refer to as *bilateralism*). In the first stage, each country simultaneously announces whether or not it wants to sign a free trade agreement (FTA) with each of its trading partners. Country i's announcement is denoted by σ_i and its strategy set Ω_i consists of four possible announcements: $\Omega_i = \{\{\phi,\phi\},\{j,\phi\},\{\phi,k\},\{j,k\}\}$, where $\{\phi,\phi\}$ denotes an announcement in favor of the status quo or no trade liberalization; $\{j,\phi\}$ in favor of an FTA with only country j; $\{\phi,k\}$ in favor of an FTA with only country k; and $\{j,k\}$ in favor of FTAs with both of them (which is equivalent to announcing in favor of multilateral free trade). This stage determines the underlying trade policy regime. Next, given the policy regime, countries impose their optimal tariffs. Finally, given trade agreements and tariffs, international trade and consumption take place.

The following trade policy regimes can emerge in the bilateralism game: (i) No agreement or the status quo $\langle\{\Phi\}\rangle$ prevails when no two announcements match or the only matching announcements are $\{\phi,\phi\}$; (ii) an FTA between countries i and j denoted by $\langle\{ij\}\rangle$ is formed iff countries i and j announce each other's name $j\epsilon\sigma_i$ and $i\epsilon\sigma_j$; (iii) two independent FTAs in which i is the common member denoted by $\langle\{ij,ik\}\rangle$ are formed iff (1) $j\epsilon\sigma_i$ and $i\epsilon\sigma_j$ and (2) $k\epsilon\sigma_i$ and $i\epsilon\sigma_k$; and (iv) free trade, denoted by $\langle\{F\}\rangle$, obtains iff all countries announce each others' names: i.e. $\sigma_i = \{j,k\}$ for all $i,j,k = a,b,c$.[12]

Before proceeding further, we clarify two expositional points. First, the regime $\langle\{ij,ik\}\rangle$ is a 'hub and spoke' trading arrangement where the 'hub' country (i.e. i) has an independent FTA with each of the two 'spoke' countries who do not have an FTA with each other. To

[12] In order to eliminate redundant announcements, we assume that each announcement costs ε (where $\varepsilon > 0$ is arbitrarily small).

simplify notation, we denote $\langle\{ij, ik\}\rangle$ as $\langle\{ih\}\rangle$ (i.e. country i is hub). Second, while changes in the underlying trade regime result from announcement deviations by countries, it proves convenient to refer directly to regime changes rather than changes in announcements. For example, when the bilateral FTA $\langle\{ij\}\rangle$ is in place, the unilateral announcement deviation of country i from $\{j, \phi\}$ to $\{\phi, \phi\}$ alters the underlying trade regime from $\langle\{ij\}\rangle$ to no agreement $\langle\{\Phi\}\rangle$ and we refer to this announcement deviation of country i as simply a deviation from $\langle\{ij\}\rangle$ to $\langle\{\Phi\}\rangle$.

Throughout the remainder of this section as well as section 4, we maintain the following assumption:[13]

Assumption 1:

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

Let country i's welfare as a function of trade regime r be denoted by $w_i(r)$ where $r \in \{\langle\{\Phi\}\rangle, \langle\{ij\}\rangle, \langle\{jk\}\rangle, \langle\{ih\}\rangle, \langle\{jh\}\rangle \text{ or } \langle\{F\}\rangle\}$ and $i, j, k = a, b, c$. Also, let $\Delta w_i(r - v)$ denote the difference between country i's welfare under trade regimes r and v: $\Delta w_i(r - v) \equiv w_i(r) - w_i(v)$.

Since Article I of GATT forbids tariff discrimination, 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 $i, j, k = a, b, c$. Country i's optimal MFN tariff is easily calculated:

$$t_i^{\phi} \equiv Arg \max w_i(\Phi) = \frac{e}{4} \tag{7}$$

If two countries form an FTA, they remove their tariffs on each other and impose their optimal external tariffs on the non-member country: under $\langle\{ij\}\rangle$ we have $t_{ij} = t_{ji} = 0$, $t_{ik} = t_i^f$ and $t_{jk} = t_j^f$. The optimal external tariff of country i on the non-member country k is given by:

$$t_i^f \equiv Arg \max w_i(ij) = \frac{e}{11} \tag{8}$$

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

Note that under symmetry, we have $t_i^\phi = t_j^\phi = t^\phi$ and $t_i^f = t_j^f = t^f$. As in Bagwell and Staiger (1997, 1999), we find that the formation of a bilateral FTA induces each member to lower its tariff on the non-member country relative to the status quo (i.e. the model exhibits tariff complementarity): $t^f < t^\phi$.[14]

Before deriving equilibrium agreements, we report a useful lemma that is easy to establish:

Lemma 1: *Under symmetry,* $\Delta w_j(ih - F) < 0 < \Delta w_i(ih - F)$.

In other words, the hub country (i) of the hub and spoke agreement $\langle\{ih\}\rangle$ is better off relative to free trade $\langle\{F\}\rangle$ while each spoke country is worse off. Note that the hub country i enjoys privileged access in both spoke countries under $\langle\{ih\}\rangle$ since neither spoke imposes a tariff on the hub whereas both impose the tariff t^f on each other. As a result of this favorable treatment, country i is strictly better off under $\langle\{ih\}\rangle$ relative to $\langle\{F\}\rangle$. To see why the spokes are worse off under $\langle\{ih\}\rangle$ relative to $\langle\{F\}\rangle$, first note that aggregate global welfare is strictly higher under $\langle\{F\}\rangle$ relative to $\langle\{ih\}\rangle$. Since the hub is strictly better off under $\langle\{ih\}\rangle$ relative to $\langle\{F\}\rangle$ and welfare of the two spoke countries is equal due to symmetry, both spokes must be worse off under $\langle\{ih\}\rangle$ relative to $\langle\{F\}\rangle$. In fact, each spoke country has an incentive to revoke its FTA with the hub and become an outsider facing an FTA between the other two countries: $\Delta w_j(ik - ih) > 0$.

We are now ready to derive equilibrium trade agreements under bilateralism. To economize space, we provide a rather condensed discussion of Nash equilibria. It is straightforward to show that given Assumption 1, *the status quo* $\langle\{\Phi\}\rangle$*, a bilateral FTA* $\langle\{ij\}\rangle$*, and free trade* $\langle\{F\}\rangle$ *are all Nash equilibria under bilateralism.* The hub and spoke regime fails to be an equilibrium because $\Delta w_j(ik - ih) > 0$. A bilateral FTA $\langle\{ij\}\rangle$ is a Nash equilibrium because

$$\Delta w_i(ij - \Phi) = \Delta w_j(ij - \Phi) > 0 \tag{9}$$

[14]See Bagwell and Staiger (1997, 1999) and Saggi and Yildiz (2009) for a detailed discussion of the tariff complementarity effect and Estevadeordal et. al. (2008) for empirical evidence in its support. It is worth noting that tariff complementarity also arises in simple general equilibrium models of trade agreements such as Bond et. al. (2004).

i.e. a member country of a bilateral FTA has no unilateral incentive to break the agreement.[15] Global free trade $\langle\{F\}\rangle$ is a Nash equilibrium because Lemma 1 implies that no country will deviate to become a spoke and the deviation to become a non-member facing an FTA is ruled out because

$$\Delta w_i(F - jk) > 0 \tag{10}$$

To deal with the multiplicity problem and to capture the process of FTA formation in a more realistic fashion, we now isolate Nash equilibria that are coalition proof or *stable* (as in Dutta and Mutuswami, 1997).[16] We begin by considering the potential stability of free trade $\langle\{F\}\rangle$. Given Lemma 1, we need to consider only two joint deviations from free trade:

(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 $\langle\{\Phi\}\rangle$.

Deviation JF1 can be ruled out because, under symmetry, no two countries have a joint incentive to exclude the third country from free trade:

$$\Delta w_i(F - ij) > 0 \tag{11}$$

Finally, it is immediate from (9) and (11) that the three countries have no incentive to deviate jointly from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$ so that JF2 is ruled out. In fact, this inequality also implies that $\langle\{\Phi\}\rangle$ is not a stable Nash equilibrium since the joint deviation of countries i and j from $\langle\{\Phi\}\rangle$ to $\langle\{ij\}\rangle$ is self-enforcing: both countries benefit from this joint deviation and it is immune to further unilateral deviations by virtue of the fact that $\langle\{ij\}\rangle$ is a Nash equilibrium. By similar logic, it is easy to see that $\langle\{ij\}\rangle$ also fails to be stable since the joint deviation of all three countries to $\langle\{F\}\rangle$ is self-enforcing. Thus, we have shown the following:

Proposition 1: *Given symmetry, free trade* $\langle\{F\}\rangle$ *is the only stable trade agreement under*

[15] Due to the tariff complementarity effect, the formation of an FTA generates a positive externality for the non-member: $\Delta w_k(ij - \Phi) > 0$.

[16] Following Bernheim et. al. (1987), a coalitional deviation is self-enforcing if a proper subset of players in the deviating coalition have no incentive to undertake a *further* deviation. Note that an alternative approach would have been to use the notion of a strong Nash equilibrium (SNE). However, the use of CPNE is more appealing since a SNE must be immune to *any* joint deviations, even those that are not self-enforcing.

bilateralism.

It is worth noting that the above result also obtains under the network formation game of Goyal and Joshi (2006) under which countries pursue only bilateral trade agreements. Interestingly, Goyal and Joshi (2006) interpret this result as establishing the compatibility of bilateralism and free trade. As we shall see below, by providing a comparison of bilateralism and multilateralism our model suggests an alternative interpretation of this result.

We now analyze a scenario where any trade liberalization undertaken by countries must be multilateral or non-discriminatory in nature. This analysis helps assess how the prospects of attaining global free trade are affected if countries cannot form bilateral FTAs that, by their very nature, discriminate against the non-member country.

4 Endogenous agreements under multilateralism

Under a multilateral approach to trade liberalization (or simply *multilateralism*), the strategy set of country i is $\Omega_i = \{\phi, M\}$, $j \neq k \neq i$. In other words, each country announces either in favor of or against multilateral trade liberalization. If all three countries announce in favor, they choose the jointly optimal set of tariffs which, in our model, are equal to zero. Thus, a multilateral agreement in which all countries participate necessarily leads to global free trade. If only two countries (say i and j) announce in favor of multilateralism, they jointly choose their optimal tariffs subject to the constraint that they must not discriminate against country k – i.e. in accordance with the MFN clause of the WTO, any tariff cut undertaken by either country i or j must apply to imports from both its trading partners. Formally, countries i and j sign the multilateral agreement $\langle\{ij^m\}\rangle$ when individual country announcements are as follows: $\sigma_i = M$, $\sigma_j = M$, $\sigma_k = \phi$. Finally, note that if two (or more) countries announce against multilateralism, the status quo $\langle\{\Phi\}\rangle$ prevails under which each country imposes its individually optimal MFN tariff on every other country.

If countries i and j agree to sign the multilateral agreement $\langle\{ij^m\}\rangle$ they choose the tariff

pair $(t_i^m,\, t_j^m)$ to solve

$$(t_i^m, t_j^m) \equiv Arg\max\left[w_i(ij^m) + w_j(ij^m)\right] \tag{12}$$

Under symmetry, $t_i^m = t_j^m = t^m$ and this jointly optimal MFN tariff t^m is given by:

$$t^m = \frac{e}{7} \text{ where } t^m < t^\phi = \frac{e}{4} \tag{13}$$

Since $t^m < t^\phi$, it is immediate that under the multilateral agreement $\langle\{ij^m\}\rangle$ countries i and j lower their tariffs on each other as well as on the non-participating country (i.e. k). This result implies that under $\langle\{ij^m\}\rangle$ country k benefits from the multilateral trade liberalization undertaken by the other two countries without having to offer any liberalization in return since it retains its individually optimal MFN tariff t^ϕ on countries i and j. Furthermore, it is worth emphasizing that country k faces lower tariffs in export markets when the other two countries implement the bilateral FTA $\langle\{ij\}\rangle$ relative to when they sign the multilateral agreement $\langle\{ij^m\}\rangle$, i.e., $t^f < t^m$. In other words, the *degree of trade liberalization* undertaken by two countries is lower when they sign the multilateral agreement $\langle\{ij^m\}\rangle$ than when they sign the bilateral agreement $\langle\{ij\}\rangle$: $t^\phi - t^m < t^\phi - t^f$. Despite this, the welfare of the non-member country (k) is higher under $\langle\{ij^m\}\rangle$ compared to $\langle\{ij\}\rangle$: $w_k(ij) < w_k(ij^m)$. This is because the non-member is subject to *discriminatory treatment* in each member's market under the bilateral FTA $\langle\{ij\}\rangle$: while countries i and j face zero tariffs in each other's market under $\langle\{ij\}\rangle$, country k faces the tariff t^f. By contrast, such discriminatory treatment is absent under $\langle\{ij^m\}\rangle$ since trade liberalization undertaken by countries i and j is extended to country k on an MFN basis. As we will show below, this fundamental difference between bilateral and multilateral trade liberalization plays a crucial role in our analysis.

As under bilateralism, it is straightforward that the status quo $\langle\{\Phi\}\rangle$ is also a Nash equilibrium under multilateralism. Furthermore, $\langle\{ij^m\}\rangle$ is not a Nash equilibrium because the outside country (k) actually benefits from joining the agreement $\langle\{ij^m\}\rangle$ thereby converting it to $\langle\{F\}\rangle$:

$$\Delta w_k(F - ij^m) > 0 \tag{14}$$

The above inequality also implies that no country has a unilateral incentive to deviate from free trade. Thus, under symmetry, free trade $\langle\{F\}\rangle$ is also a Nash equilibrium. Furthermore, since all three countries benefit from a joint deviation from the status quo to free trade (from which there are no further coalitional deviations), we have:

Proposition 2: *Given symmetry, free trade* $\langle\{F\}\rangle$ *is the only stable agreement under multilateralism.*[17]

A comparison of Propositions 1 and 2 shows that when countries are symmetric, multilateralism is sufficient to reach global free trade. One interpretation of this result is that if the move from the status quo to global free trade were to confer equal gains upon all countries (which is what happens when countries have symmetric endowments), nothing would be lost by forsaking the freedom to pursue bilateral FTAs since such agreements would *not* even arise in equilibrium. Does this imply that bilateralism is irrelevant for the ultimate objective of achieving global free trade? Or are there circumstances under which the quest for global free trade is affected in a material way by the freedom (or the lack of it) to pursue bilateral FTAs? We show next that when endowment levels are asymmetric across countries, bilateral FTAs can not only arise in equilibrium but the freedom to sign such agreements can be necessary for achieving multilateral free trade.

5 Trade liberalization under asymmetry

From hereon, we drop the assumption that endowment levels are symmetric across countries. In what follows, the size of a country is measured by its endowment of non-numeraire/protected goods relative to others. In this context, it is worth recalling that each country's endowment of the (unique) good it imports is zero and that asymmetry in endowments translates directly into asymmetries of volume of exports. In other words, an increase in a country's endowment in this model increases its exports of non-numeraire/protected goods without increasing its imports of such goods (since the model is partial equilibrium in nature and lacks any income

[17] In fact, we can show that under symmetry global free trade is the unique strong Nash equilibrium under both bilateralism and multilateralism.

effects). Indeed, since the country with the largest endowment of non-numeraire goods faces relatively smaller suppliers, its imports of such goods are the smallest. It is worth emphasizing that in our model no country is a price taker on world markets – in fact each country is the unique importer of a single good and therefore has market power that it can exploit via a tariff.

We next derive optimal tariffs under each regime under asymmetry.

5.1 Optimal tariffs

Country i's optimal non-discriminatory (or MFN) tariff is given by:

$$t_i^{\phi} \equiv Arg\max w_i(\Phi) = \frac{e_j + e_k}{8} \tag{15}$$

Note that a country's MFN tariff increases with the endowments of its trading partners. Similar to (8), when countries i and j form a bilateral FTA $\langle\{ij\}\rangle$, they abolish tariffs on each other and choose their external tariffs independently. We have

$$t_i^f \equiv Arg\max w_i(ij) = \frac{5e_k - 4e_j}{11} \text{ and } t_j^f \equiv Arg\max w_j(ij) = \frac{5e_k - 4e_i}{11} \tag{16}$$

It is easy to see that the external tariff of an FTA member increases with the endowment of the non-member whereas it decreases with that of its FTA partner.[18] Similarly, a comparison of t_i^{ϕ} and t_i^f implies that the magnitude of the tariff complementarity effect increases with the size of partner country's endowment whereas it decreases with the endowment of the non-member country. To guarantee that all tariffs are positive and non-prohibitive, given (16) we assume that $\min\{e_i, e_j, e_k\} \geq \frac{4}{5}\max\{e_i, e_j, e_k\}$.

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

$$t_i^m = \frac{2e_k - e_j}{7} \text{ and } t_j^m = \frac{2e_k - e_i}{7} \tag{17}$$

[18] It is obvious that the same optimal tariff obtains for a spoke country under a hub and spoke trading regime. By contrast, since the hub has an FTA with both spokes, it practices free trade.

5.2 Incentives for bilateral trade liberalization

How does a country's incentive to form a bilateral FTA with another depend on the distribution of endowments across countries? We address this key question by breaking it up into parts and stating three related lemmas:[19]

Lemma 2a: *Let country j be an FTA partner of country i under regime r but not under regime v and let the status of country k be the same under both regimes (i.e. either it is a partner of country i under both regimes or not). Then, the following holds:* $\frac{\partial \Delta w_i(r-v)}{\partial e_j} \leq 0 \leq \frac{\partial \Delta w_i(r-v)}{\partial e_i}$.

The intuition underlying the $\frac{\partial \Delta w_i(r-v)}{\partial e_i} \geq 0$ is as follows. Due to the smaller volume of their exports, countries with smaller endowments benefit less from tariff reductions granted by others. Similarly, such countries have relatively more to lose from eliminating their own optimal tariffs since these tariffs apply to relatively larger import volumes (or to relatively inelastic export supply curves). Thus, a country's willingness to enter into a bilateral trade agreement with another depends positively on its own endowment.

A similar intuition underlies the other inequality (i.e. $\frac{\partial \Delta w_i(r-v)}{\partial e_j} \leq 0$). The smaller the endowment of a country's partner, the larger the increase in its export surplus from the elimination of its partner's optimal tariff and the smaller the loss due to its own trade liberalization since the tariff reduction applies to a smaller volume of imports (due to the smaller size of its partner). The two inequalities reported in Lemma 2a imply that a country prefers to form a bilateral FTA with the smaller of its two trading partners:

$$w_i(ij) \geq w_i(ik) \text{ iff } e_k \geq e_j. \tag{18}$$

How does the endowment level of a competing exporter, denoted by k, affect the incentive of country i to form a bilateral FTA with country j?

Lemma 2b*: Let country j be an FTA partner of country i under regime r but not under regime v and let the status of country k be the same under both regimes (i.e. either it is a*

[19] Welfare levels under all possible regimes are reported in the appendix and these can be used to prove Lemma 2a through Lemma 4.

partner of country i under both regimes or not). Then,

(i) $\frac{\partial \Delta w_i(r-v)}{\partial e_k} \leq 0$ *if country k is an FTA partner of country j under regimes r and v;*

whereas

(ii) $\frac{\partial \Delta w_i(r-v)}{\partial e_k} \geq 0$ *if country k is not an FTA partner of country j under regimes r and v.*

The first part of the above lemma captures the idea that when country k is already an FTA partner of country j, country i's welfare gain from a bilateral FTA with country j decreases with the endowment of country k. Why is this true? Recall that both countries i and k export the same good to country j (i.e. they are competing exporters). When country k already enjoys free access to country j's market, the larger is country k's endowment the smaller the increase in country i's export surplus that results from the trade liberalization undertaken by country j. The intuition behind part (ii) of the lemma is analogous – when its rival exporter (i.e. country k) is *not* an FTA partner of country j, the strategic advantage gained by country i in country j's market from signing the bilateral FTA $\langle\{ij\}\rangle$ increases in country k's size, making the FTA more valuable from its perspective.

The following lemma examines the welfare implications of being a hub country (say i under $\langle\{ih\}\rangle$) relative to other trade regimes:

Lemma 3: $w_i(ih) > \max\{w_i(ij),\, w_i(F),\, w_i(\Phi)\}$ *for all* $i, j = a, b, c$.

The fact that hub country prefers $\langle\{ih\}\rangle$ to $\langle\{F\}\rangle$ informs us that the first part of Lemma 1 generalizes to the case of asymmetric endowments. The intuition is the same as that under symmetry: relative to free trade, the hub country enjoys privileged access in both spoke countries. One implication of this lemma is worth stressing: a hub country has no incentive to unilaterally revoke either one or both of its FTAs.

5.3 Multilateral trade liberalization

How do the incentives of a country to form (or join to) a multilateral agreement depend on the underlying endowment structure?

Lemma 4: *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 all of the inequalities reported in Lemma 4 is quite analogous to that which underlies parallel results under bilateralism with only one exception – i.e. whereas $\frac{\partial \Delta w_i(ij-\phi)}{\partial e_k} > 0$ under bilateralism when country k is a non-member country, 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 multilateral agreement $\langle\{ij\}\rangle^m$, countries i and j lower their tariffs on not only to each other but also on country k whereas under the bilateral agreement $\langle\{ij\}\rangle$ they only lower tariffs on each other. The larger is country k's endowment, the smaller the increase in the export surplus that countries i and j obtain due to the multilateral agreement $\langle\{ij^m\}\rangle$ since their rival exporter (i.e. country k) captures a larger share of their markets.

5.4 Equilibrium trade agreements under asymmetry

To highlight the crucial role played by asymmetry, it proves instructive to consider a scenario where two countries (denoted by l and l') have larger endowments than the third (denoted by s; refereed to as the 'smaller' country).[20] Accordingly, let the pattern of endowment asymmetry be given by:[21]

Assumption 2a:

$$e_s = \frac{e}{\theta} < e_l = e_{l'} = e \text{ and } 1 \leq \theta \leq \frac{5}{4} \tag{19}$$

To avoid redundancy, we focus directly on stable agreements under bilateralism. First consider the perspective of the two larger countries. We know from Lemma 1 that spoke countries are worse off relative to free trade under symmetry. Similarly, Lemma 2a and Lemma 2b imply that $\frac{\partial \Delta w_l(F-l'h)}{\partial e_s} \leq 0$ and $\frac{\partial \Delta w_l(F-sh)}{\partial e_s} \leq 0$. Thus, a larger country (say l) under free

[20] As noted earlier, in our model no country is 'small' in the traditional sense since all three can influence their terms of trade. Hence we use the word 'smaller' as opposed to 'small'.

[21] An earlier version of the paper also analyzed the case where $e_s = e_{s'} = \frac{e}{\theta} < e_l = e$ and $\frac{5}{4} \geq \theta > 1$. To save space, we omit a detailed discussion here and simply note that analogous results hold under this alternative type of asymmetry.

trade has no incentive to revoke one of its FTAs and become a spoke:

$$\Delta w_l(F - sh) > 0 \text{ and } \Delta w_l(F - l'h) > 0 \text{ for all } \theta \tag{20}$$

Similarly, we know from (10) that under symmetry, starting from global free trade a country has no incentive to unilaterally revoke its two FTAs. Lemma 2a and Lemma 2b reinforce this result for the larger countries under asymmetry. We have:

$$\frac{\partial \Delta w_l(F - sl')}{\partial e_s} = \underbrace{\frac{\partial \Delta w_l(F - sh)}{\partial e_s}}_{\leq 0} + \underbrace{\frac{\partial \Delta w_l(sh - sl')}{\partial e_s}}_{\leq 0} \leq 0 \tag{21}$$

Therefore, a larger country (say l) prefers $\langle\{F\}\rangle$ to $\langle\{sl'\}\rangle$:

$$\Delta w_l(F - sl') > 0 \text{ for all } \theta \tag{22}$$

Thus, inequalities (20) and (22) show that a larger country has no unilateral incentive to defect from free trade.

It is immediate from (20) that a joint defection from free trade to any hub and spoke regime does not occur. This implies we only need to consider three possible joint defections from free trade:

(JF1): Joint deviation of l and s from $\langle\{F\}\rangle$ to $\langle\{sl\}\rangle$.

(JF2): Joint deviation of l and l' from $\langle\{F\}\rangle$ to $\langle\{ll'\}\rangle$.

(JF3): Joint deviation of l and s or l and l' or all countries from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$.

We know from inequality (11) that under symmetry ($\theta = 1$) no two countries benefit from excluding the third country from free trade. Furthermore, we show in the appendix that $\Delta w_l(F - sl)$ is monotonically decreasing in θ and that $\Delta w_l(F - sl) > 0$ at the smallest possible endowment of country s (i.e. at $\theta = \frac{5}{4}$) . This implies that joint deviation JF1 cannot occur: $\Delta w_l(F - sl) > 0$ for all θ. It is immediate from this and inequality (18) that the two larger countries have no incentive to jointly deviate from $\langle\{F\}\rangle$ to $\langle\{ll'\}\rangle$: $\Delta w_l(F - ll') > 0$ for all

θ. Therefore, joint deviation JF2 is also ruled out. Finally, the fact that $\Delta w_l(F - sl) > 0$ together with inequalities (9) and Lemma 2a implies that joint deviation JF3 also does not occur: $\Delta w_l(F - \Phi) > 0$ for all θ. Thus, we have shown the following:

Lemma 5: *Suppose Assumption 2a holds. Then, there exist no unilateral or coalitional deviations of larger countries from free trade.*

Lemma 5 suggests that the stability of global free trade depends critically upon the preferences of the smaller country. Let $\theta_i(r - v)$ denote the critical threshold at which country i is indifferent between regimes r and v. Direct calculations yield

$$\Delta w_s(F - ll') \geq 0 \text{ iff } \theta \leq \theta_s(F - ll') \text{ and } \Delta w_s(F - lh) \geq 0 \text{ iff } \theta \leq \theta_s(F - lh) \qquad (23)$$

where we show in the appendix that $\theta_s(F - ll') < \theta_s(F - lh)$. Together with (23), this implies that free trade is stable iff $\theta \leq \theta_s(F - ll')$. What happens when $\theta > \theta_s(F - ll')$? Parts (ii) and (iii) of the following proposition (proved in the appendix) addresses this question:

Proposition 3a: *Given Assumption 2a, the following hold under bilateralism:*

(i) $\langle\{F\}\rangle$ *is uniquely stable when* $\theta \leq \theta_s(F - ll')$*;*

(ii) both $\langle\{sl\}\rangle$ *and* $\langle\{ll'\}\rangle$ *are stable when* $\theta_s(F - ll') \leq \theta \leq \theta_{l'}(lh - sl)$*; and*

(iii) $\langle\{ll'\}\rangle$ *is uniquely stable when* $\theta \geq \theta_s(F - ll')$*.*

– Figure 1here –

Proposition 3a 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. This implies that Proposition 1 does not require symmetry but rather that the degree of endowment asymmetry be sufficiently small. Part (ii) says that if the degree of endowment asymmetry is moderate, both a bilateral trade agreement between a smaller and a larger country and a bilateral FTA between the two larger countries are stable whereas part (iii) says that if the degree of endowment asymmetry is sufficiently large, only an FTA between the two larger countries is stable – in such a situation, the smaller country

prefers being a non-member to participating in multilateral free trade.[22]

It is noteworthy that multiple stable equilibria obtain when the degree of endowment asymmetry is moderate – i.e. when $\theta_s(F - ll') \leq \theta \leq \theta_{l'}(lh - sl)$. Since theory offers no guidance about which of these equilibria might be observed, we examine both of these possibilities hereafter.

6 When, why, and how bilateralism matters

To see how the ability to form bilateral FTAs matters, suppose countries were to follow only a multilateral approach to trade liberalization. Under such an approach, there are only stable equilibria: $\langle\{sl^m\}\rangle$ and $\left\langle\{ll'^m\}\right\rangle$. To see why, first note that Lemma 5 implies that there can be no coalitional deviations from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$. Furthermore, Lemma 4 implies that the larger country l' has no incentive to unilaterally deviate from $\langle\{F\}\rangle$ to $\langle\{sl^m\}\rangle$. This implies that $\langle\{sl^m\}\rangle$ is not stable. In fact, the only deviation from free trade that we need to consider is the unilateral deviation of the smaller country from $\langle\{F\}\rangle$ to $\left\langle\{ll'^m\}\right\rangle$. It turns out that this deviation does not occur if the degree of endowment asymmetry is small enough:

$$\Delta w_s(F - ll'^m) \geq 0 \text{ iff } \theta \leq \theta_s(F - ll'^m) \tag{24}$$

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 that $\Delta w_l(ll'^m - \Phi) > 0$ under symmetry ($\theta = 1$). Since $\frac{\partial \Delta w_l(ll'^m - \phi)}{\partial e_s} < 0$ (Lemma 4) we have

$$\Delta w_l(ll'^m - \Phi) > 0 \text{ for all } \theta \tag{25}$$

Inequalities (24) and (25) imply that the multilateral agreement $\left\langle\{ll'^m\}\right\rangle$ is stable when $\theta > \theta_s(F - ll'^m)$.

[22] A similar result obtains in Ludema (2002) where asymmetry takes the form of transportation costs as opposed to endowments.

Proposition 3b: *Given Assumption 2a,* $\langle\{F\}\rangle$ *is stable when* $\theta \leq \theta_s(F - ll'^m)$. *Otherwise* $\left\langle\{ll'^m\}\right\rangle$ *is stable.*

Figure 2 shows stable agreements under multilateralism.

– Figure 2 here–

Recall that under bilateralism, global free trade is stable only when $\theta \leq \theta_s(F - ll')$ whereas it is stable under multilateralism only when $\theta \leq \theta_s(F - ll'^m)$. A straightforward comparison of these critical thresholds delivers one of our major results:

Proposition 4: *Given Assumption 2a, the following hold:*

(i) $\theta_s(F - ll'^m) < \theta_s(F - ll')$; *and*

(ii) over the parameter range $\theta_s(F - ll'^m) < \theta \leq \theta_s(F - ll')$ *the unique stable agreement under bilateralism is* $\langle\{F\}\rangle$ *whereas under multilateralism it is* $\left\langle\{ll'^m\}\right\rangle$.

– Figure 3 here–

Part (i) of proposition 4 says that free trade is stable over a *larger* parameter space when countries are free to sign bilateral FTAs relative to when they cannot. Part (ii) demonstrates that there exist circumstances where the freedom to pursue bilateral FTAs is *necessary* for achieving global free trade. This happens because the smaller country has a greater incentive to choose global free trade under bilateralism due to the fact that it is discriminated against its rival exporter in each larger member country's market under the bilateral FTA $\langle\{ll'\}\rangle$ whereas it suffers no such disadvantage under the multilateral agreement $\langle\{ll'^m\}\rangle$ – i.e. opting out of global free trade is relatively costlier for the smaller country under bilateralism. It is noteworthy that this result obtains even though the smaller country faces lower tariffs in its export markets under the bilateral FTA between the two larger countries $\langle\{ll'\}\rangle$ relative to that under the multilateral agreement $\left\langle\{ll'^m\}\right\rangle$. Thus, if the degree of endowment asymmetry is not too large, the threat of a bilateral FTA between the two larger countries and the discrimination that is inherent to such a trade agreement can be necessary to nudge the smaller country to

announce in favor of global free trade. The very fact that the multilateral agreement is non-discriminatory in nature makes it *less* effective in altering the trade-off facing the smaller country since it does not lose as much from opting out of global free trade.[23]

When $\theta > \theta_s(F - ll')$, global free trade fails to obtain under both bilateralism and multilateralism. Intuitively, if the smaller country is sufficiently small, then even the possibility of a bilateral FTA between the two larger countries is not enough to induce it to opt for global free trade. When such is the case, one possible way forward is to ask how global welfare compares under the equilibrium agreements that obtain under bilateralism and multilateralism. From Proposition 3b we know that under bilateralism both $\langle\{ll'\}\rangle$ and $\langle\{sl\}\rangle$ are stable agreements when $\theta_s(F - ll') < \theta \leq \theta_{l'}(lh - sl)$ whereas $\langle\{ll'\}\rangle$ is uniquely stable when $\theta > \theta_{l'}(lh - sl)$. Furthermore, Proposition 3b says that when global free trade does not obtain under multilateralism, $\left\langle\{ll'^m\}\right\rangle$ emerges as the unique stable equilibrium. Thus, when $\theta > \theta_s(F - ll')$, we need to consider two possible scenarios: (1) $\langle\{ll'\}\rangle$ is stable or (2) $\langle\{sl\}\rangle$ is stable. First, consider scenario (1) and note that *lower* internal and external tariffs (thus freer trade) obtain under $\langle\{ll'\}\rangle$ relative to $\left\langle\{ll'^m\}\right\rangle$: $t_l^m > t_l^f$. Thus, larger trade volumes and higher aggregate world welfare obtain under $\langle\{ll'\}\rangle$ relative to $\left\langle\{ll'^m\}\right\rangle$: $\Delta ww(ll' - ll'^m) > 0$.

Now consider scenario (2) where $\langle\{sl\}\rangle$ is the stable bilateral agreement. We show in the appendix that $\Delta ww(sl - ll'^m) > 0$ when $\theta_s(F - ll') < \theta < \theta_{l'}(lh - sl)$. In other words, over the relevant parameter range, global welfare is higher under the bilateral agreement $\langle\{sl\}\rangle$ relative to the multilateral agreement $\left\langle\{ll'^m\}\right\rangle$. Thus, when free trade is out of reach, the option to pursue bilateral FTAs can yield welfare-improving trade liberalization that is foregone under a strictly multilateral approach.[24]

---

[23] It is worth noting here that the result reported in part (i) of Proposition 4 depends crucially on the fact that, relative to the status quo, the degree of the positive externality enjoyed by the non-member is stronger under the multilateral agreement $\left\langle\{ll'^m\}\right\rangle$ relative to the bilateral FTA $\langle\{ll'\}\rangle$. In an alternative model, the relative strength of this positive externality under the two types of agreements might be reversed. And if so, the incentives for participating in global free trade could be stronger when discriminatory bilateral agreements are not permitted.

[24] Of course, aggregate world welfare does not necessarily speak to the fate of individual countries. It is immediate form (24) that, when the freedom to pursue bilateral FTAs is *necessary* for achieving global free trade, large countries are better off under bilateralism relative to multilateralism whereas the small country's fate is the opposite.

7 Political economy considerations

In order to determine whether and how the presence of political economy considerations affects our main results, suppose countries put additional weight on producer surplus relative to tariff revenue and consumer surplus:

$$w_i = \sum_z CS_i^z + TR_i + (1+\eta)\sum_z PS_i^z \tag{26}$$

Since the model lacks an import competing industry, the additional weight on producer surplus (i.e. η) has no effect on the tariffs countries implement under no agreement and a bilateral FTA. As a result, domestic surplus stays unchanged under the bilateralism game. However, by increasing the importance of the export surplus gain that results from trade liberalization undertaken by partner countries, the additional weight on producer surplus makes participation in FTAs more desirable.

On the other hand, under the multilateral agreement $\langle\{ij^m\}\rangle$ countries i and j internalize each other's additional weights on export profits and their respective MFN tariffs fall with η:

$$t_i^m = \frac{2e_k - (1+3\eta)e_j}{7} \text{ and } t_j^m = \frac{2e_k - (1+3\eta)e_i}{7} \tag{27}$$

As before, to guarantee that all tariffs are non-negative, we assume that (i) $\min\{e_i, e_j, e_k\} \geq \frac{4}{5}\max\{e_i, e_j, e_k\}$ and (ii) $\eta \leq \frac{1}{5}$. For the case of symmetric endowments, we prove the following result in the appendix:

Proposition 5: *Given symmetric endowments ($e_i = e$ for all i), the following hold:*

(i) $\langle\{F\}\rangle$ is uniquely stable under bilateralism for all η whereas

(ii) $\langle\{F\}\rangle$ is uniquely stable under multilateralism only when $\eta \leq \frac{1}{10}$. Otherwise, $\langle\{ij^m\}\rangle$ is the stable equilibrium.

Thus, provided that the degree of political economy pressure is not too small (i.e. $\eta > \frac{1}{10}$), the freedom to pursue bilateral FTAs can be *necessary* for achieving global free trade *even when countries have symmetric endowments*. To understand the intuition behind this result, consider

the perspective of the outside country (i.e. country k) under $\langle\{ij^m\}\rangle$: under multilateralism, the higher the weight on producer surplus, the lower the tariffs faced by the outside country in its export markets. As a result, the incentive to opt out of multilateral trade liberalization increases with the degree of political economy pressure η. By contrast, such a result does not obtain when the agreement involved is a bilateral FTA since the external tariffs of FTA members are independent of the degree of political economy pressure faced by them.

Now suppose that countries have asymmetric endowments. Let the pattern of endowment asymmetry be given by Assumption 2 and two countries (denoted by l and l') have larger endowments than the third (denoted by s). As before, let $\theta_i^p(r-v)$ denote the critical threshold at which country i is indifferent between regimes r and v, where $\theta_i^p(r-v)$ is now a function of η.

It is straightforward to show that Lemma 5 continues to hold under bilateralism – i.e. the two larger countries do not deviate either unilaterally or coalitionally from free trade. Thus, once again, the stability of free trade depends critically upon the unilateral preferences of the smaller country. Direct calculations establish that:

$$\Delta w_s(F-ll') \geq 0 \text{ iff } \theta \leq \theta_s^p(F-ll') \text{ and } \Delta w_s(F-lh) \geq 0 \text{ iff } \theta \leq \theta_s^p(F-lh) \qquad (28)$$

where $\theta_s^p(F-ll') < \theta_s^p(F-lh)$ for all η. Therefore, free trade is stable under bilateralism iff $\theta \leq \theta_s^p(F-ll')$.

Similarly, under multilateralism, while the larger countries have no incentive to opt out of free trade, the smaller country benefits from deviating unilaterally from $\langle\{F\}\rangle$ to $\left\langle\{ll'^m\}\right\rangle$ if the degree of endowment asymmetry is large enough: $\Delta w_s(F-ll'^m) \geq 0$ iff $\theta \leq \theta_s^p(F-ll'^m)$. It follows that free trade is stable under multilateralism when $\theta \leq \theta_s^p(F-ll'^m)$. We prove the following result in the appendix:

Proposition 6: *Given Assumption 2, the following hold:*

(i) $\theta_s^p(F-ll'^m) < \theta_s^p(F-ll')$ *for all* η*;*

(ii) over the parameter range $\theta_s^p(F-ll'^m) < \theta \leq \theta_s^p(F-ll')$*, bilateralism yields* $\langle\{F\}\rangle$ *as*

the stable equilibrium whereas multilateralism yields $\left\langle\{ll'^{m}\}\right\rangle$;

(iii) when $\theta > \theta_s^p(F - ll')$*, multilateralism yields* $\left\langle\{ll'^{m}\}\right\rangle$ *as the stable equilibrium whereas bilateralism yields* $\langle\{ll'\}\rangle$ *and also* $\langle\{sl\}\rangle$ *when in addition* $\theta \leq \theta_{l'}^p(lh - sl)$ *and*

(iv) world welfare is higher under bilateralism relative to multilateralism for all η*.*

Proposition 6 provides a confirmation of our key insight that a country that is reluctant to liberalize has a greater incentive to opt out of global free trade under multilateralism relative to bilateralism. Moreover, this insight receives even stronger support as the additional weight on producer surplus increases since under multilateralism the smaller country does not suffer from discrimination (as it does under an FTA) and the tariff it faces as an outsider decreases with η.

Figure 4 illustrates equilibrium agreements in the (η, θ) space to show how the degree of political economy pressure and endowment asymmetry jointly determine equilibrium outcomes under bilateralism and multilateralism:

– Figure 4 here –

Note from Figure 4 that free trade is more likely to be stable under bilateralism as η increases. By contrast, free trade fails to be stable under multilateralism when η is sufficiently large ($\eta > \frac{1}{10}$). Finally, when $\theta > \theta_s^p(F - ll')$, free trade fails to obtain under both bilateralism and multilateralism. Under such a case, stable agreements under bilateralism ($\langle\{ll'\}\rangle$ and $\langle\{sl\}\rangle$) lead to higher global welfare than the one under multilateralism ($\left\langle\{ll'^{m}\}\right\rangle$).

8 Conclusion

One of the striking features of today's global policy landscape is the widespread prevalence of preferential trade agreements. Only a handful of countries are not involved in one and most simultaneously participate in several such agreements. Jagdish Bhagwati (1991) famously raised concern about the potential adverse effects of the pursuit of preferential trade agreements on the prospects of multilateral trade liberalization. His work led to a rich body of research

that has illuminated various aspects of the multi-faceted relationship between preferential and multilateral trade liberalization. However, this literature has often tended to treat bilateral trade agreements as exogenous or only considered an endogenous trade agreement between a pair of countries while treating the third country as a silent observer. By contrast, we present a model in which all countries are free to pursue both bilateral and multilateral agreements. To determine whether bilateralism hampers or facilitates the obtainment of global free trade, we also derive stable equilibria under a purely multilateral approach to trade agreements. This analysis helps shed light on the pros and cons of bilateralism and multilateralism.

A central result of this paper is that bilateralism can actually provide an impetus to multilateral trade liberalization. The point is that a country that is choosing whether or not to participate in global free trade must consider its fate under the agreement that would emerge in the *absence* of its participation. Due to the fact that a bilateral trade agreement discriminates against the outsider whereas a multilateral agreement does not, a non participating country is worse off under the former relative to the latter. As a result, a country's incentive to opt for free trade is stronger when the alternative to free trade is a bilateral agreement between the other two countries as opposed to a multilateral one.

An important implication of our analysis is that to properly account for the role of bilateralism, we need to better understand *why* countries choose to enter into bilateral agreements when multilateral trade liberalization is an option. To this end, the model suggests that the debate regarding preferential versus multilateral liberalization is moot in the absence of some type of asymmetry across countries. This is because, in our model, whether or not countries are free to pursue bilateral trade agreements, global free trade is the only stable equilibrium under symmetry. This result demonstrates that heterogeneity across countries with respect to the benefits that they enjoy from global free trade may be a critical determinant of the success of a purely multilateral approach to trade liberalization. In our view, such heterogeneity has received insufficient attention in the literature and its role merits further research.

9 Appendix

This appendix contains supporting calculations and brief proofs of propositions stated in the paper.

Supporting calculations

We begin by reporting welfare levels under different policy regimes. Under bilateralism, we have:

$$w_i(F) = \frac{[(\sum_j \frac{e_j}{3})^2 + \sum_j (\frac{e_j}{3})^2]}{2} + e_i(2\alpha - \frac{2e_i + \sum_{j\neq i} e_j}{3})$$

$$w_i(\Phi) = \frac{\sum_{j\neq i}[\frac{3(e_i+e_j)}{8}]^2}{2} + (\frac{e_j+e_k}{4})^2 + e_i(2\alpha - \frac{6e_i + \sum_{j\neq i} 3e_j}{8})$$

$$w_i(ij) = \frac{[\frac{3(e_i+e_j)}{8}]^2 + [\frac{(2e_k+5e_i)}{11}]^2}{2} + \frac{(4e_k^2 + 3e_j^2 - 2e_je_k)}{22} + e_i[2\alpha - (\frac{73e_i}{88} + \frac{3e_j}{8} + \frac{2e_k}{11})]$$

$$w_i(jk) = \frac{\sum_{j\neq i}(\frac{7e_i+e_j}{11})^2}{2} + (\frac{e_j+e_k}{4})^2 + e_i(2\alpha - \frac{14e_i + \sum_{j\neq i} e_j}{11})$$

$$w_i(ih) = \frac{1}{2}\left[[\frac{(e_j+e_k)}{3}]^2 + \sum_{j\neq i}[\frac{(2e_j+5e_i)}{11}]^2\right] + e_i(2\alpha - \frac{10e_i + \sum_{j\neq i} 2e_j}{11})$$

and

$$w_i(jh) = \frac{(\frac{e_i+e_k}{3})^2 + (\frac{7e_i+e_j}{11})^2}{2} + \frac{(4e_k^2 + 3e_j^2 - 2e_je_k)}{22} + e_i(2\alpha - \frac{32e_i + 11e_k + 3e_j}{33})$$

Under multilateralism, we have:

$$w_i(ij^m) = 2\alpha e_i + \frac{3(e_i+e_j)(3e_j - 13e_i)}{128} + \frac{(2e_i+3e_k)(3e_k - 12e_i) + (3e_j+e_k)(e_j+5e_k)}{98}$$

$$w_i(jk^m) = 2\alpha e_i + (\frac{e_j+e_k}{4})^2 + \frac{(3e_i+2e_j)(2e_j - 11e_i) + (3e_i+2e_k)(2e_k - 11e_i)}{98}$$

Welfare levels under symmetry can be calculated by setting each country's endowment to e in

the formulae above. The relevant comparisons under symmetry are as follows:

$$\begin{aligned}
\Delta w_i(ij-\Phi) &= \frac{47}{2}(\frac{e}{44})^2>0;\ \Delta w_k(ij-\Phi)=23(\frac{e}{44})^2>0;\\
\Delta w_i(ih-F) &= 23(\frac{e}{33})^2>0;\ \Delta w_j(F-ih)=\frac{29}{2}(\frac{e}{33})^2>0;\\
\Delta w_i(ih-ij) &= \frac{1039}{2}(\frac{e}{132})^2>0;\ \Delta w_j(ik-ih)=\frac{161}{2}(\frac{e}{132})^2>0;\\
\Delta w_i(F-jk) &= \frac{13}{3}(\frac{e}{22})^2>0;\Delta w_i(F-ij)=\frac{101}{6}(\frac{e}{22})^2>0
\end{aligned}$$

and

$$\Delta w_i(ij^m-\Phi)=\frac{1}{14}(\frac{e}{4})^2>0;\ \Delta w_k(F-ij^m)=\frac{1}{3}(\frac{e}{14})^2>0$$

Proof of Lemma 3

First consider part (i). We know from Lemma 1 that $\Delta w_i(ih-F)>0$ under symmetry. Next, note that $\frac{\partial \Delta w_i(ih-F)}{\partial e_i}=\frac{134(e_j+e_k)-320e_i}{33^2}<0$, $\frac{\partial \Delta w_i(ih-F)}{\partial e_j}=\frac{134e_i-85e_j}{33^2}>0$ and $\frac{\partial \Delta w_i(ih-F)}{\partial ek}=\frac{134e_i-85e_k}{33^2}>0$. At $e_i=\frac{4e}{5}$ and $e_j=e_k=e$, we have $\Delta w_i(ih-F)=3(\frac{e}{11})^2>0$. Using analogous arguments, we can establish parts (ii) and (iii).

Critical thresholds

Since $e_l=e_{l'}$, we must have $\theta(F-lh)=\theta(F-l'h)$. Furthermore, $\theta(F-ll')_s\cong 1.0398$ and $\theta(F-lh)_s\cong 1.1487$.

Inequalities from the text

We have

$$\Delta w_l(F-sl)\mid_{\theta=\frac{5}{4}}=\frac{7}{2}(\frac{e}{12})^2>0 \text{ and } \frac{\partial \Delta w_l(F-sl)}{\partial \theta}=-\frac{1357\theta-1211}{6336\theta^3}<0$$

Proof of Proposition 3a

Note from (9) that under symmetry two countries always benefit from forming a bilateral FTA. Also, we know from Lemma 2b that $\frac{\partial \Delta w_l(ll'-\phi)}{\partial e_s}>0$. Next, note that $\Delta w_l(ll'-\Phi)\mid_{\theta=\frac{5}{4}}=\frac{3}{10}(\frac{e}{8})^2>0$. This implies that $\langle\{\Phi\}\rangle$ is not stable:

$$\Delta w_l(ll'-\Phi)>0 \text{ for all } \theta \tag{29}$$

Furthermore, inequalities (18) and (29) together imply that

$$\Delta w_l(sl - \Phi) > 0 \text{ for all } \theta \tag{30}$$

Consider now the smaller country's perspective under $\langle\{sl\}\rangle$. From Lemma 2a, we know that $\frac{\partial \Delta w_l(sl-\phi)}{\partial e_s} > 0$. Further note that $\Delta w_s(sl - \Phi) \mid_{\theta=\frac{5}{4}} = \frac{719}{2}(\frac{e}{440})^2 > 0$. This implies

$$\Delta w_s(sl - \Phi) > 0 \text{ for all } \theta \tag{31}$$

It is immediate from (20) that the two larger countries jointly defect from $\langle\{sh\}\rangle$ to $\langle\{F\}\rangle$ and this defection is self-enforcing since neither has an incentive to further defect (Lemma 5). Thus, $\langle\{sh\}\rangle$ is not stable. Similarly, the smaller country defects unilaterally from $\langle\{lh\}\rangle$ to $\langle\{ll'\}\rangle$ so that $\langle\{lh\}\rangle$ is not stable.[25]

Next, we provide conditions under which $\langle\{sl\}\rangle$ and $\langle\{ll'\}\rangle$ are stable. There exist five possible coalitional deviations from $\langle\{sl\}\rangle$:

(JSL1): Deviation of l and l' from $\langle\{sl\}\rangle$ to $\langle\{ll'\}\rangle$.

(JSL2): Deviation of s and l' from $\langle\{sl\}\rangle$ to $\langle\{sh\}\rangle$.

(JSL3): Deviation of l and l' from $\langle\{sl\}\rangle$ to $\langle\{lh\}\rangle$.

(JSL4): Deviation of all countries from $\langle\{sl\}\rangle$ to $\langle\{l'h\}\rangle$.

(JSL5): Deviation of all countries from $\langle\{sl\}\rangle$ to $\langle\{F\}\rangle$.

Note from (18) that country l will not defect from $\langle\{sl\}\rangle$ to $\langle\{ll'\}\rangle$. Thus, JSL1 is ruled out. Next consider JSL2 and JSL3. We know from Lemma 3 that country s (l) has an incentive to defect from $\langle\{sl\}\rangle$ to $\langle\{sh\}\rangle$ ($\langle\{lh\}\rangle$). For these deviations to occur, the choice of country $l^{'}$ is pivotal. We have

$$\Delta w_{l'}(sh - sl) \geq 0 \text{ iff } \theta \geq \theta_{l'}(sh - sl) = 1.0639 \tag{32}$$

and

$$\Delta w_{l'}(lh - sl) \geq 0 \text{ iff } \theta \geq \theta_{l'}(lh - sl) = 1.0629 \tag{33}$$

[25] An analogous discussion applies to $\langle\{l'h\}\rangle$.

Since $\theta \geq \theta_{l'}(lh - sl) > \theta_{l'}(sh - sl)$, JSL3 is the binding deviation.

Now consider JSL4. Since the smaller country has an incentive to unilaterally deviate from $\langle\{l'h\}\rangle$ to $\langle\{ll'\}\rangle$, even if JSL4 occurs, it is not self-enforcing. Finally, we know from (23) that JSL5 occurs when $\theta < \theta_s(F - lh)$ and it is self enforcing only if $\theta < \theta_s(F - ll^{'})$. Thus, $\langle\{sl\}\rangle$ is stable iff $\theta_{l'}(lh - sl) \geq \theta \geq \theta_s(F - ll')$.

We now derive conditions under which $\langle\{ll'\}\rangle$ is stable. Inequality (29) implies that there can be no deviations from $\langle\{ll'\}\rangle$ to $\langle\{\Phi\}\rangle$. Now consider the following coalitional deviations:

(JLL1): Deviation of s and l from $\langle\{ll'\}\rangle$ to $\langle\{sl\}\rangle$.

(JLL2): Deviation of s and l from $\langle\{ll'\}\rangle$ to $\langle\{lh\}\rangle$.

(JLL3): Deviation of all countries from $\langle\{ll'\}\rangle$ to $\langle\{sh\}\rangle$.

(JLL4): Deviation of all countries from $\langle\{ll'\}\rangle$ to $\langle\{F\}\rangle$.

From Lemma 3, it is immediate that JLL1 is not a self-enforcing deviation since country l has an incentive to further deviate to $\langle\{lh\}\rangle$. Moreover, JLL2 is ruled out since country s does not have an incentive to defect from $\langle\{ll'\}\rangle$ to $\langle\{lh\}\rangle$. We also know from (20) that even if JLL3 occurs, the two larger countries further deviate from $\langle\{sh\}\rangle$ to $\langle\{F\}\rangle$. Thus the initial deviation is not self-enforcing. Finally, (23) implies that all countries deviate from $\langle\{ll'\}\rangle$ to $\langle\{F\}\rangle$ when $\theta < \theta_s(F - ll')$ and this deviation is self-enforcing. Thus, $\langle\{ll'\}\rangle$ is stable iff $\theta \geq \theta_s(F - ll')$.

Other inequalities from the text

We have $\Delta w_s(F - ll'^m) \geq 0$ iff $\theta \leq \theta_s(F - ll'^m) = 1.0149$. Furthermore,

$$\frac{\partial \Delta ww(sl - ll'^m)}{\partial \theta} = -\frac{8605\theta + 1949}{2\theta}(\frac{e}{308\theta})^2 < 0$$

Note that when $\theta = \theta_{l'}(lh - sl)$, $\Delta ww(sl - ll'^m) > 0$. Thus, $\Delta ww(sl - ll'^m) > 0$ when $\theta_s(F - ll') \leq \theta < \theta_{l'}(lh - sl)$.

Other calculations

$$\Delta w_s(F - \Phi) = (\frac{e}{24\theta})^2 \frac{51 - 25\theta^2 - 2\theta}{2} > 0 \text{ for all } \theta$$

$$\Delta w_s(F - ss') \leq 0 \text{ iff } \theta \geq \theta_s(F - ss') = 1.0845$$

$$\Delta w_s(F - s'l) \leq 0 \text{ iff } \theta \geq \theta_s(F - s'l) = 1.0810$$

$$\Delta w_s(F - s'h) \leq 0 \text{ iff } \theta \geq \theta(F - s'h)_s = 1.1814$$

Proof of Proposition 5

$$\begin{aligned}
\Delta w_i(F - \Phi) &= \frac{e^2(8\eta + 1)}{48} > 0; \Delta w_i(F - jk) = \frac{176\eta + 13}{3}(\frac{e}{22})^2 > 0; \\
\Delta w_i(F - ij) &= \frac{616\eta + 101}{6}(\frac{e}{44})^2 > 0; \Delta w_i(F - jh) = \frac{132\eta + 29}{2}(\frac{e}{33})^2 > 0 \\
\Delta w_i(F - jk) &= \frac{(18\eta + 1)(1 - 10\eta)}{3}(\frac{e}{14})^2 \geq 0 \text{ iff } \eta \leq \frac{1}{10}
\end{aligned}$$

Proof of Proposition 6

Part (i) is immediate from the fact that $\theta_s^p(F - ll') < \theta_s^p(F - ll'^m)$ for all feasible η.

Parts (ii) and (iii): The following is immediate from the proof proposition 3a

$$\Delta w_l(ll' - \Phi) > 0, \Delta w_l(sl - \Phi) > 0, \text{ and } \Delta w_s(sl - \Phi) > 0 \text{ for all } \theta, \eta \tag{34}$$

These inequalities imply that $\langle\{\Phi\}\rangle$ is not stable.

We know from (20) that two larger countries jointly defect from $\langle\{sh\}\rangle$ to $\langle\{F\}\rangle$ when $\eta = 0$. This defection also occurs for any $\eta > 0$ and is self-enforcing since neither country has an incentive to further defect. Therefore, $\langle\{sh\}\rangle$ is not stable.

Now consider $\langle\{lh\}\rangle$.[26] The smaller country defects from $\langle\{lh\}\rangle$ to $\langle\{ll'\}\rangle$ unless $\theta \leq \theta_s^p(lh - ll')$ where $\theta_s^p(lh - ll')$ is defined by $\Delta w_s(lh - ll') = 0$. However, when $\theta \leq \theta_s^p(lh - ll')$, the joint deviation of countries s and l' from $\langle\{lh\}\rangle$ to $\langle\{F\}\rangle$ is self-enforcing so that $\langle\{lh\}\rangle$ fails to be stable.

Next consider $\langle\{sl\}\rangle$ and $\langle\{ll'\}\rangle$. We know from (34) that unilateral defections from $\langle\{sl\}\rangle$ to $\langle\{\Phi\}\rangle$ and $\langle\{ll'\}\rangle$ to $\langle\{\Phi\}\rangle$ do not occur. There exist five possible coalitional deviations from $\langle\{sl\}\rangle$:

[26] An analogous discussion applies to $\langle\{l'h\}\rangle$.

(JSL1-P): Deviation of l and l' from $\langle\{sl\}\rangle$ to $\langle\{ll'\}\rangle$.

(JSL2-P): Deviation of s and l' from $\langle\{sl\}\rangle$ to $\langle\{sh\}\rangle$.

(JSL3-P): Deviation of l and l' from $\langle\{sl\}\rangle$ to $\langle\{lh\}\rangle$.

(JSL4-P): Deviation of all countries from $\langle\{sl\}\rangle$ to $\langle\{l'h\}\rangle$.

(JSL5-P): Deviation of all countries from $\langle\{sl\}\rangle$ to $\langle\{F\}\rangle$.

JSL1-P is ruled out since country l does not defect from $\langle\{sl\}\rangle$ to $\langle\{ll'\}\rangle$. Next consider JSL2-P and JSL3-P. We know from Lemma 3 that country s (l) has an incentive to defect from $\langle\{sl\}\rangle$ to $\langle\{sh\}\rangle$ ($\langle\{lh\}\rangle$). For these deviations to occur, the choice of country $l^{'}$ is pivotal. We have $\Delta w_{l'}(sh-sl) \geq 0$ iff $\theta \geq \theta^p_{l'}(sh-sl)$ and $\Delta w_{l'}(lh-sl) \geq 0$ iff $\theta \geq \theta^p_{l'}(lh-sl)$. Since $\theta_{l'}(lh-sl) \geq \theta_{l'}(sh-sl)$ JSL3-P is the binding deviation.

Next note that since country l has no incentive to deviate from $\langle\{sl\}\rangle$ to $\langle\{l'h\}\rangle$, JSL4-P does not occur. Finally, note that the two larger countries have an incentive to jointly deviate from $\langle\{sl\}\rangle$ to $\langle\{F\}\rangle$ whereas the smaller country defects only when it has a relatively symmetric endowment $\Delta w_s(F-sl) \geq 0$ iff $\theta \leq \theta^p_s(F-sl)$ and and this joint deviation is self enforcing only if $\theta < \theta^p_s(F-ll^{'})$. Thus, $\langle\{sl\}\rangle$ is stable iff $\theta^p_s(F-ll') \leq \theta \leq \theta^p_{l'}(lh-sl)$.

For $\langle\{ll'\}\rangle$ to be stable, we need to consider the following coalitional deviations:

(JLL1-P): Deviation of s and l from $\langle\{ll'\}\rangle$ to $\langle\{sl\}\rangle$.

(JLL2-P): Deviation of s and l from $\langle\{ll'\}\rangle$ to $\langle\{lh\}\rangle$.

(JLL3-P): Deviation of all countries from $\langle\{ll'\}\rangle$ to $\langle\{sh\}\rangle$.

(JLL4-P): Deviation of all countries from $\langle\{ll'\}\rangle$ to $\langle\{F\}\rangle$.

Note that JLL1-P is not a self-enforcing deviation since country l has an incentive to further deviate to $\langle\{lh\}\rangle$ and JLL2-P happens only when $\theta < \theta^p_s(lh-ll')$ and that it is a self-enforcing deviation. Even if JLL3-P occurs, the larger countries further deviate from $\langle\{sh\}\rangle$ to $\langle\{F\}\rangle$, making the initial deviation not self-enforcing. Finally, all countries deviate from $\langle\{ll'\}\rangle$ to $\langle\{F\}\rangle$ when $\theta < \theta^p_s(F-ll')$ and this deviation is self-enforcing. Since $\theta^p_s(lh-ll') < \theta^p_s(F-ll')$, $\langle\{ll'\}\rangle$ is stable iff $\theta \geq \theta^p_s(F-ll')$.

Under multilateralism, straightforward calculations show that (i) $\Delta w_l(ll'^m-\Phi)$; (ii) $\Delta w_l(F-\Phi) > 0$; (iii) $\Delta w_{l'}(F-sl^m) > 0$; and (iv) $\Delta w_s(F-ll'^m) \geq 0$ iff $\theta \leq \theta^p_s(F-ll'^m)$.

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, 1997. "Multilateral Tariff Cooperation During the Formation of Free Trade Areas." International Economic Review 38, 291-319.

[3] Bagwell, K., Staiger, R.W., 1999. Regionalism and multilateral tariff cooperation. In John Pigott and Alan Woodland, eds., International Trade Policy and the Pacific Rim, London: Macmillan.

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

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

[6] Bhagwati, Jagdish, Arvind Panagariya, and Pravin Krishna, eds., Trading Blocs, 1990, The MIT Press, Cambridge, MA.

[7] Bond, Eric W., Constantinos, Syropoulos, and L. Alan Winters, 2001. "Deepening of Regional Integration and Multilateral Trade Agreements," Journal of International Economics 53, 335-362.

[8] 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.

[9] Dutta, Bhaskar and Suresh Mutuswami, 1997. "Stable Networks." Journal of Economic Theory 76, 322-344.

[10] Estevadeordal, Antoni, Caroline Freund and Emanuel Ornelas, 2008. "Does Regionalism Affect Trade Liberalization Toward Non-members?" Quarterly Journal of Economics 123, 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 and Hideo Konishi, 2007. "Free Trade Networks." Journal of International Economics 72, 310-335.

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

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

[15] Jackson, Matthew and Asher Wolinsky, 1996. "A Strategic Model of Social and Economic Networks." Journal of Economic Theory 71, 44-74.

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

[17] Krugman, Paul R. "The Move Toward Free Trade Zones." in Policy Implications of Trade and Currency Zones: A Symposium Sponsored by the Federal Reserve Bank of Kansas City, Federal Reserve Bank of Kansas City, Kansas City, 7-41, 1991.

[18] Levy, Philip I., 1997. "A Political-Economic Analysis of Free Trade Agreements." American Economic Review 87, 506-519.

[19] Ludema, Rodney, 2002. "Increasing Returns, Multinationals and Geography of Preferential Trade Agreements." Journal of International Economics 56, 329-358.

[20] Ornelas, Emanuel, 2005a. "Endogenous Free Trade Agreements and the Multilateral Trading System." Journal of International Economics 67, 471-497.

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

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

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

[24] Saggi, Kamal and Halis M. Yildiz, 2006. "Bilateral Trade Agreements and the Feasibility of Multilateral Free Trade." Mimeo.

[25] Saggi, Kamal and Halis M. Yildiz, 2009. "Optimal Tariffs of Preferential Trade Agreements and the Tariff Complementarity Effect." Indian Growth and Development Review 2(1), 5-17.

[26] World Bank, 2005. Global Economic Prospects: Trade, Regionalism, and Development, The World Bank, Washington, D.C.

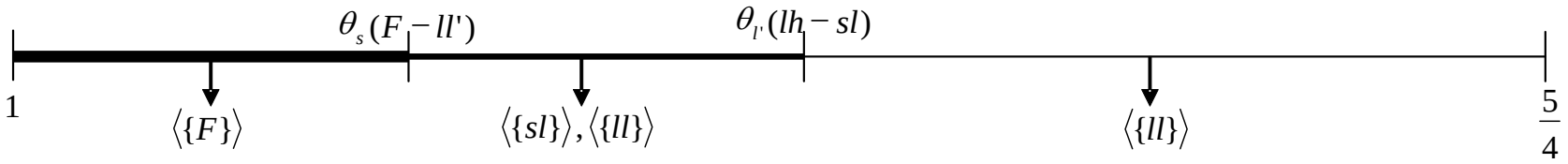

Figure 1: Stable agreements under bilateralism

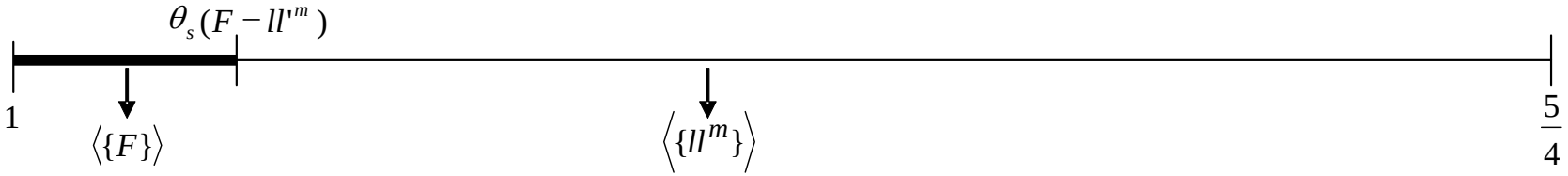

Figure 2: Stable agreements under multilateralism

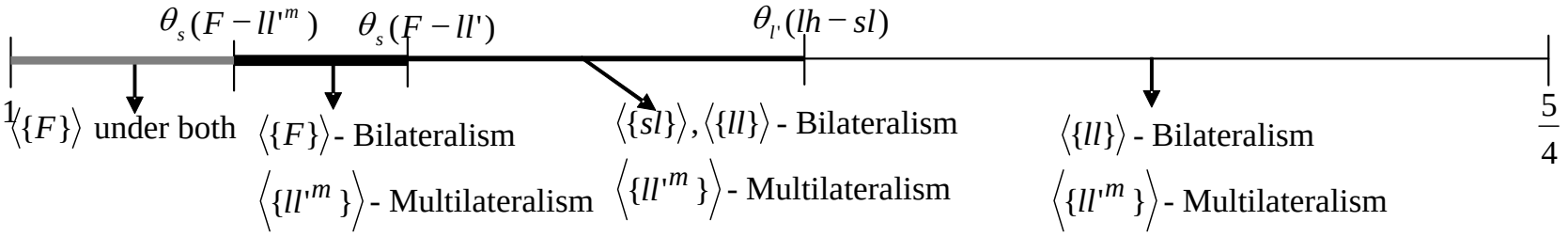

Figure 3: Bilateralism versus multilateralism

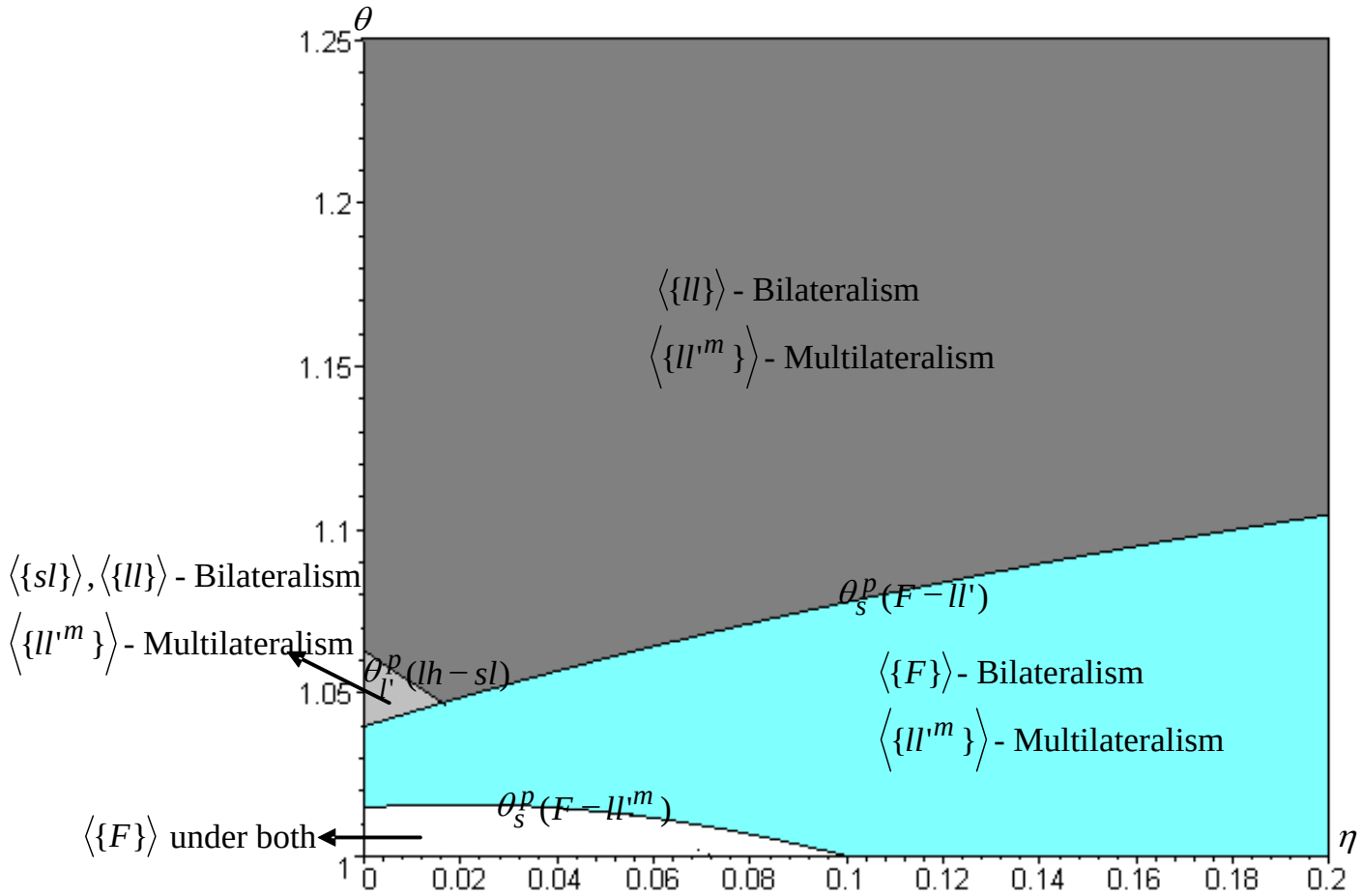

Figure 4: Political Economy considerations and bilateralism versus multilateralism