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

Kamal Saggi* and Halis Murat Yildiz†

Abstract

This paper develops an equilibrium theory of trade agreements and evaluates the relative merits of bilateralism and multilateralism. We derive coalition proof (stable) Nash equilibria of a three-country game in which each country is free to negotiate a trade agreement with only one of its trade partners, or both of them (i.e. practice free trade), or none of them (i.e. opt for the status quo). To determine whether and how bilateralism matters, we also analyze this game under the assumption that countries follow a *purely* multilateral approach to trade liberalization. Results show that: (1) under symmetry, global free trade is a stable equilibrium regardless of whether countries can pursue bilateral agreements or not; (2) when countries have asymmetric endowment levels, there exist circumstances under which free trade is a stable equilibrium only if countries are free to sign bilateral agreements; (3) welfare improving bilateral agreements can be stable when global free trade is not; and (4) while bilateralism can sometimes undermine global free trade – such as when two similar sized countries are better off under a bilateral agreement relative to global free trade – the parameter space over which this happens is relatively small and (5) the option to form bilateral trade agreements necessarily reduces the likelihood of obtaining the status quo as a stable equilibrium.

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

*Department of Economics, Southern Methodist University, Dallas, TX 75275-0496. Phone: 214-768-3274; fax: 214-768-1821; e-mail: ksaggi@smu.edu.

†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.

1 Introduction

Global trade liberalization occurs through a variety of channels, not all of which appear to be in harmony with one another. While 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 a subset of WTO members to pursue PTAs under which they can grant tariff (and other trade policy) concessions to each other that they do *not* have to extend to others.[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 free trade agreements (FTAs) and then use it to evaluate the pros and cons of bilateral and *purely* multilateral approaches to trade liberalization.

An important feature of our approach is that we allow countries to form multiple FTAs. Formally, 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 proceeds as follows. In the first stage, each country announces whether or not it wants to form an FTA with each of its trading partners. An FTA between two countries requires them to abolish tariffs on each other and it arises iff they both announce each other's name. Similarly, global free trade emerges iff all countries call each other's names. Next, given the world trade regime, international trade

[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.

and consumption take place. After analyzing equilibrium trade agreements, we examine how the adoption of a purely multilateral approach affects the likelihood of obtaining global free trade.

We show that when countries are symmetric with respect to their endowment levels, while there exist multiple Nash equilibria, global free trade is the only stable equilibrium *regardless* of whether countries can pursue bilateral agreements or not. This irrelevance result points to the importance of allowing for heterogeneity across countries. To this end, we then consider a scenario where endowment levels are unequal across countries. This analysis delivers what is perhaps our most novel result: there exist circumstances where global free trade is a stable equilibrium *only if* countries can form bilateral FTAs. The insight behind this result is as follows. Under pure multilateralism, a single country has the ability to preserve the status quo by voting against global free trade. However, if bilateral agreements are feasible, a country that chooses not to participate in any trade agreement can find itself *worse off* relative to the status quo if its trading partners implement a bilateral agreement between themselves since such preferential liberalization adversely affects its exports to their markets.[2] Anticipating this outcome, it can then become a willing participant in multilateral free trade. In fact, in our model, such an effect can *only* obtain when countries are asymmetric with respect to their endowment levels.

Our model also isolate circumstances where bilateralism undermines global free trade. For example, when both free trade and a bilateral FTA between two large countries are stable equilibria, a strictly multilateral approach can ensure that global free trade is uniquely stable. In general, when two countries are better off under global free trade relative to no agreement but have an incentive to exclude the third country, a purely multilateral approach is preferable to one that permits both bilateral and multilateral agreements. However, two points are worth noting. First, this negative effect of bilateralism obtains over a fairly small parameter space. Second, the option to form bilateral agreements necessarily lowers the likelihood of

[2] See Chang and Winters (2002) for detailed evidence regarding the adverse effects of the Latin American customs union MERCOSUR on the exports of non-member countries to MERCOSUR.

obtaining the status quo as a stable equilibrium.

Overall, our results suggest that heterogeneity across countries is 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.[3] 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. While our model considers differences in endowment levels across countries, we conjecture that international differences in technology, underlying institutions, and political economy forces should deliver similar results.[4]

In a recent paper, Aghion et. al. (2004) provide a comparison of sequential and multilateral bargaining of FTAs. While we consider similar issues, there are important differences between their approach and ours. First, in our model, *all* countries are free to negotiate FTAs and are not required to choose between joining a single grand coalition with a leading country or staying out. Second, we allow countries to form multiple bilateral FTAs. Third, unlike them we do not allow transfers between different coalitions. This is important because when transfers are possible and global free trade maximizes aggregate welfare, it emerges as the equilibrium under both sequential and multilateral bargaining.[5] When free trade does not maximize aggregate welfare, Aghion et. al. (2004) show 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.

Our approach is also related to that of Riezman (1999) who also asks whether bilateralism facilitates or hinders the achievement of global free

[3] 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.

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

[5] We obtain a similar result in the absence of transfers when endowment levels are symmetric across countries.

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. As noted above, endowment asymmetry across countries plays a crucial in determining the welfare implications of bilateralism in our model.

The relationship between preferential and multilateral liberalization, to which Bhagwati (1991) first drew attention, has frequently been analyzed in the literature in models of repeated interaction between countries – see Bagwell and Staiger (1997), Bond and Syropoulos (1996), Freund (2000), and Saggi (2006).[6] We add value to this literature by treating both bilateral and multilateral liberalization as endogenous.

2 Underlying trade model

To focus on endogenous trade agreements among asymmetric countries, we utilize an appropriately adapted version of the partial equilibrium framework developed by Bagwell and Staiger (1997 and 1998). There are three countries: a, b, and c denoted by $j = \{a, b, c\}$ and three (non-numeraire) goods: A, B, and C denoted by $i = \{A, B, C\}$. Each country's market is served by two competing exporters and J denotes the good that corresponds to the upper case value of j. For example, if $j = a$, then $J = A$. Country a is endowed with $x/2$ units of goods B and C and zero units of good A; country b is endowed with $y/2$ units of goods A and C and zero units of good B; and country c is endowed with $z/2$ units of goods A and B and zero units of good C.[7] Without loss of generality, let $x \leq y \leq z$.

The demand for good i in country j is given by

$$D(P_i^j) = \alpha - \beta P_i^j \tag{1}$$

As is well known, the above demand functions can be derived from a utility

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

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

function of the form $U(c_i) = u(c_i) + w$ where c_i denotes consumption of good i; w denotes the numeraire good; and $u(c_i)$ is quadratic and additively separable in each of the three goods. Since each country possesses only two goods while it demands all three, country J must import good j in order to consume it and it can import it from either trading partner. The pattern of trade is as follows: each country imports a single good from the other two and exports different goods to each of them. 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 be the tariff imposed by country j on its imports of good J. Ruling out prohibitive tariffs yields the following no-arbitrage conditions for good A:

$$P_A^a = P_A^b + t = P_A^c + t \tag{2}$$

Similar no-arbitrage conditions hold for the other two goods. Let M_A^a be country a's imports of good A. Since country a has no endowment of good A, we have

$$M_A^a = \alpha - \beta P_A^a \tag{3}$$

Country b's exports of good A equal its endowment of that good minus its local consumption:

$$X_A^b = y/2 - [\alpha - \beta P_A^b] \tag{4}$$

Similarly, country c's exports of good A to country a equal:

$$X_A^c = z/2 - [\alpha - \beta P_A^c] \tag{5}$$

Market clearing for good A requires that country a's imports equal the total exports of countries b and c:

$$M_A^a = X_A^b + X_A^c \tag{6}$$

Conditions analogous to (2) through (6) hold for the other two goods.

Equation (2) through (6) imply that

$$\alpha - \beta P_A^a = y/2 - [\alpha - \beta P_A^b] + z/2 - [\alpha - \beta P_A^c] \tag{7}$$

Using the non-arbitrage condition (2), we have

$$\alpha - \beta P_A^a = y/2 - [\alpha - \beta (P_A^a - t_A^b)] + z/2 - [\alpha - \beta (P_A^a - t_A^c)] \tag{8}$$

which yields the equilibrium price of good A in country a:

$$P_A^a = \frac{6\alpha - y - z}{6\beta} + \frac{2t}{3} \tag{9}$$

As expected, the price of good A in country a increases in its tariff and decreases in the endowment levels of the other two countries. From equations (2) and (9) we have,

$$P_A^b = P_A^c = \frac{6\alpha - y - z}{6\beta} - \frac{t}{3} \tag{10}$$

The positive effect of a country's tariff t on its terms of trade is evident from equations (9) and (10): a small increase in t lowers the price collected by foreign exporters. In fact, exactly 1/3rd of the total tariff t is paid by foreign exporters with the rest of the burden falling on the domestic economy.

Using these prices, the volume of trade is easily calculated. Country a's imports of good A equal:

$$M_A^a = \frac{y + z}{6} - \frac{2\beta t}{3} \tag{11}$$

where the exports of each of its trading partners are

$$X_A^b = \frac{2y - z}{6} - \frac{\beta t}{3} \text{ and } X_A^c = \frac{2z - y}{6} - \frac{\beta t}{3} \tag{12}$$

As expected, country a's total imports decrease with its tariff.

By design the model examines country j's trade protection towards only good J (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 $(y + z)/6$ units of good A) whereas country c faces the lowest volume of imports of such goods (it imports $(x + y)/6$ units of good C).[8] Note also that country j's imports of good J do *not* equal its exports of other non-numeraire goods. For example, under free trade, country a exports $(2x - z)/6$ units of good B to country b and $(2x - y)/6$ units of good C to country c and these are lower than its

[8] The same ranking applies with respect to the value of imports $P_J^j M_J^j$ so long as $\alpha > x + y + 2z$, which is a minor condition that is assumed to hold. Under this condition, value of country a's imports (of good A) exceeds the value of country b's imports (of good B) and country c's value of imports (of good C).

imports of good A from either country: $0 < 2x - z < 2x - y < y + z$. 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. Thus, for example, country a's welfare is given by:

$$W^a(t) = \sum_J CS_J^a + \sum_J PS_J^a + TR^a \tag{13}$$

where country a's total consumer surplus equals

$$\sum_J CS_J^a = \frac{1}{2\beta}\left[(M_A^a)^2 + (\alpha - \beta P_B^a)^2 + (\alpha - \beta P_C^a)^2\right] \tag{14}$$

its producer surplus equals

$$\sum_J PS_J^a = (x/2 - X_B^a)\, P_B^a + (x/2 - X_C^a)\, P_C^a + X_B^a(P_B^b - t) + X_C^a(P_C^c - t) \tag{15}$$

and its tariff revenue is given by

$$TR^a = t(X_A^b + X_A^c) \tag{16}$$

Aggregate world welfare is defined as the sum of each country's welfare. We now describe the process of FTA formation.

3 Endogenous trade agreements

Under the status quo, each country imposes a non-discriminatory tariff t on both its trading partners. If two countries form an FTA, they remove their tariffs on each other while retaining them on the non-member country. The process of FTA formation is as follows. Each country simultaneously announces whether or not it wants to form an FTA with each of its trading partners (country i's announcement is denoted by α_i). Country i's strategy set Ω_i consists of four possible announcements:

$$\Omega_i = \{\{\phi, \phi\}, \{j, \phi\}, \{\phi, k\}, \{j, k\}\} \tag{17}$$

where the announcement $\{\phi, \phi\}$ by country i is in favor of the status quo; $\{j, \phi\}$ is in favor of an FTA with only country j; $\{\phi, k\}$ is in favor of an FTA with only country k; and $\{j, k\}\}$ is in favor of FTAs with both of them (which is equivalent to country i announcing in favor of free trade).

The following policy regimes can emerge in this game: (i) Status quo or no agreement $\langle\{\Phi\}\rangle$ prevails when no two announcements match or when everyone announces $\{\phi, \phi\}$; (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 $j \epsilon \alpha_i$ and $i \epsilon \alpha_j$; (iii) two independent bilateral FTAs in which i is the common member denoted by $\langle\{ij, ik\}\rangle$ are formed iff (1) $j \epsilon \alpha_i$ and $i \epsilon \alpha_j$ and (2) $k \epsilon \alpha_i$ and $i \epsilon \alpha_k$; and (iv) free trade, denoted by $\langle\{F\}\rangle$, obtains iff all countries announce each others' names: i.e. $\alpha_i = \{j, k\}$ for all $i, j, k = a, b, c$. Given the policy regime determined by the process of FTA formation, international trade and consumption take place.

It is worth noting here that the regime under which there exist two independent bilateral FTAs (i.e. $\langle\{ij, ik\}\rangle$) is a 'hub and spoke' trading arrangement where the common member (i.e. country i) is the hub while each of the other two countries is a spoke. To economize notation, denote a hub and spoke arrangement with country i as the hub by $\langle\{ih\}\rangle$.

Before deriving the equilibrium trade agreements, we clarify an expositional point: while changes in the underlying trade regime result from announcement deviations by countries, it proves more 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 $\langle\{\Phi\}\rangle$.

We now derive equilibrium trade agreements.

4 Equilibrium analysis in the benchmark model

Throughout this section and the next, we maintain the following assumption:[9]

[9] All supporting calculations are contained in the appendix.

Assumption 1 (Symmetric endowments): $x = y = z = e$.

In order to guarantee market access for all exporters, we exclude prohibitive tariff levels and assume:

$$t < \overline{t} = \frac{e}{4\beta} \tag{18}$$

4.1 Nash equilibria

It is straightforward that the status quo $\langle\{\Phi\}\rangle$ is always a Nash equilibrium since no country has an incentive to announce another's name if the latter does not announce its name in return. Is a bilateral FTA $\langle\{ij\}\rangle$ a Nash equilibrium too? Let country i's welfare as a function of trade regime r be denoted by $w_i(r)$ where $r \in \{\langle\{\Phi\}\rangle, \langle\{ij\}\rangle, \langle\{ij, ik\}\rangle$, or $\langle\{F\}\rangle\}$ and $i, j, k = a, b, c$. It is easy to show that

$$w_i(ij) \geq w_i(\Phi) \tag{19}$$

i.e. a member country of a bilateral FTA has no unilateral incentive to break the agreement. Thus, a bilateral FTA is indeed a Nash equilibrium.

Now consider a hub and spoke arrangement $\langle\{ih\}\rangle$ as a candidate for Nash equilibrium. It is easily shown that the country i's welfare under $\langle\{ih\}\rangle$ is higher than that under $\langle\{\Phi\}\rangle$:

$$w_i(ih) \geq w_i(\Phi) \tag{20}$$

Thus, the hub country (i) under $\langle\{ih\}\rangle$ has no unilateral incentive to revoke its two FTAs. Furthermore, the hub country (i) has no incentive to unilaterally break one of its agreements:

$$w_i(ih) \geq w_i(ij) \tag{21}$$

Does a spoke country have a unilateral incentive to cancel its FTA with the hub? If a spoke country does revoke its FTA with the hub, it ends up becoming an outsider facing an FTA between the other two countries and it is easy to show that

$$w_j(ih) \geq w_j(ik) \text{ iff } t \leq t_h \equiv \frac{e}{7\beta} \tag{22}$$

This implies that a hub and spoke arrangement $\langle\{ih\}\rangle$ is a Nash equilibrium so long as the tariff rate t is sufficiently low ($t \leq t_h$). An noteworthy aspect of a hub and spoke arrangement is the following:

Lemma 1: *Under symmetry, the hub country (i) of* $\langle\{ih\}\rangle$ *is better off relative to free trade while each spoke country is worse off:*

$$w_j(ih) = w_k(ih) \leq w_j(F) = w_k(F) \leq w_i(ih) \tag{23}$$

Intuitively, a hub country enjoys privileged access in both foreign countries while it itself practises free trade. Thus, its export surplus is higher than that under free trade while its domestic surplus is no different.

The only remaining candidate for a Nash equilibrium is global free trade $\langle\{F\}\rangle$. For free trade to be a Nash equilibrium, we need to rule out the following two representative deviations of country k:

- UF1: From $\langle\{F\}\rangle$ to $\langle\{ih\}\rangle$ (or $\langle\{jh\}\rangle$).
- UF2: From $\langle\{F\}\rangle$ to $\langle\{ij\}\rangle$.

It is obvious from Lemma 1 that UF1 cannot occur. Also, country k has no unilateral incentive to revoke both of its FTAs since it is worse off as a non-member than it is under free trade:

$$w_i(F) \geq w_i(jk) \tag{24}$$

We have shown the following:

Proposition 1: *No agreement* $\langle\{\Phi\}\rangle$*, a bilateral FTA* $\langle\{ij\}\rangle$*, and free trade* $\langle\{F\}\rangle$ *are all Nash equilibria. In addition, a hub and spoke arrangement such as* $\langle\{ih\}\rangle$ *is also a Nash equilibrium iff* $t \leq t_h$*.*

To deal with the multiplicity of equilibria described in Proposition 1 and to capture the process of FTA formation in a more realistic fashion, we now isolate Nash equilibria that are coalition proof (i.e. are immune to credible or self-enforcing coalitional deviations). Following Dutta and Mutuswami's (1997) terminology, we refer to coalition proof Nash equilibria as *stable* equilibria.

4.2 Stable equilibria

Proposition 1 notes that four policy regimes are Nash equilibria. Which, if any, are stable? We begin with no agreement $\langle\{\Phi\}\rangle$. Consider a member country's welfare under $\langle\{ij\}\rangle$ relative to $\langle\{\Phi\}\rangle$. It is immediate from (19) that countries i and j have an incentive to jointly deviate from $\langle\{\Phi\}\rangle$ to $\langle\{ij\}\rangle$. Since this joint deviation is self-enforcing, $\langle\{\Phi\}\rangle$ is *not* stable.

Now consider a hub and spoke arrangement $\langle\{ih\}\rangle$. We know from inequality (23) that countries j and k benefit from a joint deviation from $\langle\{ih\}\rangle$ to $\langle\{F\}\rangle$. Once again, this joint deviation is self-enforcing since $\langle\{F\}\rangle$ is a Nash equilibrium. As a result, $\langle\{ih\}\rangle$ is also *not* stable.

Two candidates for stable equilibria remain: $\langle\{F\}\rangle$ and $\langle\{ij\}\rangle$. For $\langle\{F\}\rangle$ to be stable, we need to rule out three types of coalitional deviations:

- JF1: Deviation of i and j from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$.
- JF2: Deviation of j and k from $\langle\{F\}\rangle$ to $\langle\{ij, ik\}\rangle$.
- JF3: Deviation of i and j from $\langle\{F\}\rangle$ to $\langle\{ij\}\rangle$.

It is straightforward to show that each country is better off under global free trade relative to the status quo so that joint deviation JF1 cannot occur. Similarly, Lemma 1 implies that deviation JF2 cannot occur. Can deviation JF3 be ruled out? We can show that

$$w_i(F) \geq w_i(ij) \text{ iff } t \geq t_l \equiv \frac{e}{9\beta} \tag{25}$$

Thus, if the tariff t is sufficiently large, country i is worse off under $\langle\{ij\}\rangle$ relative to $\langle\{F\}\rangle$ and deviation JF3 cannot occur. Under such a scenario, $\langle\{F\}\rangle$ is immune to all coalitional (as well as unilateral) deviations and is a stable equilibrium.[10] But what if countries i and j have an incentive to jointly deviate from free trade to $\langle\{ij\}\rangle$ (i.e. $t < t_l$ holds)? To determine whether this deviation is self-enforcing or not, we need to consider two further deviations from $\langle\{ij\}\rangle$:

[10] In fact, since free trade is immune to even those coalitional deviations that are not self-enforcing, it is actually a strong Nash equilibrium under symmetry.

- FD1: Deviation of country i from $\langle\{ij\}\rangle$ to $\langle\{\Phi\}\rangle$.
- FD2: Deviation of country i from $\langle\{ij\}\rangle$ to $\langle\{ih\}\rangle$.

It is immediate from inequality (19) (i.e. $w_i(ij) > w_i(\Phi)$) that deviation FD1 cannot occur. What about deviation FD2? It is easy to show that country i indeed has an incentive to further deviate from $\langle\{ij\}\rangle$ to the pair of bilateral FTAs $\langle\{ih\}\rangle$:

$$w_i(ih) \geq w_i(ij) \tag{26}$$

Since deviation FD2 will indeed occur, the initial joint deviation of countries i and j from $\langle\{F\}\rangle$ to $\langle\{ij\}\rangle$ (i.e. deviation JF3) is *not* self-enforcing. Thus, $\langle\{F\}\rangle$ is stable even when $t < t_l$. Can $\langle\{ij\}\rangle$ also be stable when $t < t_l$? Note from (22) that over this range of tariffs, countries i and k indeed have a joint incentive to deviate from $\langle\{ij\}\rangle$ to $\langle\{ih\}\rangle$. We know from Proposition 2 that $\langle\{ih\}\rangle$ is a Nash equilibrium when $t < t_l$. Therefore, the initial deviation of countries i and k from $\langle\{ij\}\rangle$ to $\langle\{ih\}\rangle$ is self-enforcing. As a result, $\langle\{ij\}\rangle$ is *not* stable.

We have demonstrated the following:

Proposition 2A: *When countries are free to pursue both bilateral and multilateral trade agreements, free trade is the unique stable equilibrium under symmetry.*

We now analyze a scenario where countries follow a purely multilateral approach to trade liberalization.

4.3 The irrelevance of bilateral agreements

Under a purely multilateral approach, the strategy set of country i is $\Omega_i = \{\{\phi, \phi\}, \{j, k\}\}$, $j \neq k \neq i$. In other words, each country can now announce either in favor of or against free trade – i.e. any agreement must include everyone. If all countries announce in favor, the outcome is free trade. If not, the status quo prevails.

Under pure multilateralism, *any* unilateral or joint deviation from global free trade results in no agreement. But since each country prefers global free

trade to status quo under symmetry, no unilateral or joint deviations from free trade can occur:

Proposition 2B: *Under symmetry, free trade is the unique stable equilibrium even under pure multilateralism.*

Thus when countries are symmetric, a purely multilateral approach to trade liberalization yields exactly the same outcome as an approach under which both bilateral and multilateral agreements are possible. In other words, if global trade liberalization were to confer equal gains upon all countries, nothing would be lost by forsaking the freedom to pursue bilateral agreements since such agreements would *not* even arise in equilibrium.

Given this result, it is natural to ask: under what circumstances. if any, does the freedom to pursue bilateral agreements actually matter? We show next that such indeed is the case when countries have unequal endowment levels.

5 When, why, and how bilateralism matters?

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. This is useful because the volume of a country's exports (of non-numeraire goods) is positively related to its endowment while the volume of its imports of such goods is negatively related to it. This implies that countries with smaller endowments have relatively more to gain from using tariffs. Similarly, due to the smaller volume of their exports, such countries have less to lose from other countries' tariffs. Thus, a country's willingness to enter into a bilateral trade agreement with another depends positively on its own endowment.

We should note here 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 can be exploited via a tariff. Thus, the traditional notion of a 'small' country – i.e. one that cannot influence its terms of trade – does not apply here.

It proves instructive to focus on the following cases: (i) two countries have bigger endowments than the third and (ii) two countries have smaller

endowments than the third.

5.1 One small and two large countries

Let country b and c's endowment levels exceed that of country a:

Assumption 2: $x < y = z = e$.

To exclude prohibitive tariff levels, we assume $x > x^{cr} = \frac{e}{2} + 2\beta t$.

5.1.1 Equilibrium trade agreements

To avoid redundancy, we focus directly on stable equilibria. First consider the perspective of the two large countries (i.e. b and c).

Lemma 2: *There exists no self-enforcing deviations of large countries from free trade.*

Lemma 2 is important because it demonstrates that, starting at free trade, the two large countries *cannot* exclude the small country by forming a bilateral FTA between themselves. Even if they were to jointly deviate from free trade to a bilateral FTA, each has an incentive to further deviate to position itself as a hub by forming an independent agreement with the excluded country. This incentive for further deviation renders the original joint deviation from free trade non-credible. Lemma 2 is noteworthy because one of the policy concerns with respect to the proliferation of preferential trade agreements has been that such agreements may serve as devices for excluding smaller countries from the multilateral trading system. What this result shows is that, at least in our model, large countries are not the source of this problem.

On the other hand, there exists a critical threshold endowment level (e^{ϕ}) such that opening up its market is unattractive to the small country when its endowment falls below it:

$$w_a(\Phi) \geq w_a(F) \text{ iff } x \leq e^{\phi} \equiv \frac{2e - 3\beta t}{2} \tag{27}$$

This is because what country a gives up in terms of domestic surplus when it eliminates its tariff is not adequately offset by what it gains in export markets.

Given these results, it is clear that the viability of global free trade depends critically upon the preferences of the small country. Following the definition of e^{ϕ} (in 27), we can define threshold endowment levels e^{bc} and e^{bh} above which country a prefers free trade to $\langle\{bc\}\rangle$ and $\langle\{bh\}\rangle$ respectively:[11]

$$w_a(bh) \geq w_a(F) \text{ iff } x \leq e^{bh} \equiv \frac{3e - 7\beta t}{4} \tag{28}$$

and

$$w_a(bc) \geq w_a(F) \text{ iff } x \leq e^{bc} \equiv \frac{3e}{4} \tag{29}$$

From (27), (28) and (29) it immediately follows that $e^{bh} < e^{bc} < e^{\phi}$ since $x > x^{cr}$. We are now ready to characterize equilibrium trade agreements. The following proposition is proved in the appendix:

Proposition 3: *Let* $x_t \equiv e^{bc} + \frac{7\beta t}{4}$. *Free trade is stable if* $e^{bc} \leq x$ *whereas a bilateral trade agreement between the two large countries is stable if either (a)* $x \leq e^{bc}$ *or (b)* $e^{bc} < x < x_t$ *and* $t < t_l$. *Finally, none of the other trade agreements are stable.*

– Figure 1: Stable agreements –

Proposition 3 shows that multiple stable equilibria obtain when $t < t_l$ and $e^{bc} < x < x_t$. To understand the source of this multiplicity, note that over this parameter space, countries a and b have an incentive to jointly deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. However, this joint deviation is not self enforcing since country b has an incentive to further deviate from $\langle\{ab\}\rangle$ to $\langle\{bh\}\rangle$. As a result, $\langle\{F\}\rangle$ is stable. On the other hand, over the same parameter space, all countries have an incentive to jointly deviate from $\langle\{bc\}\rangle$ to $\langle\{F\}\rangle$ but this joint deviation is also not self enforcing since countries a and b have a joint incentive to further deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. As a result, $\langle\{bc\}\rangle$ is also stable. Since theory offers no guidance about which of these equilibria might be observed, to avoid redundancy we focus on the case where $\langle\{F\}\rangle$ is stable when $t < t_l$ and $e^{bc} < x < x_t$, and then briefly comment on the scenario where $\langle\{bc\}\rangle$ is stable.

[11] Since $y = z = e$ it must be that $e^{bh} = e^{ch}$.

5.1.2 How bilateralism facilitates global free trade

To see how bilateralism matters, suppose countries were to follow a purely multilateral approach to trade agreements. Under such an approach, there can only be two possible equilibrium outcomes: $\langle\{\Phi\}\rangle$ and $\langle\{F\}\rangle$. It is straightforward to show that $\langle\{F\}\rangle$ obtains as a stable agreement when $x > e^{\phi}$ whereas $\langle\{\Phi\}\rangle$ obtains otherwise (see Figure 2 for an illustration). Intuitively, when the small country's endowment falls below the critical threshold e^{ϕ}, it prefers status quo to free trade since the worsening of its terms of trade implied by the removal of its tariff dominates the relatively small increase in its export profits that results from the elimination of foreign tariffs (recall that level of a country's exports are proportional to its endowment).

– Figure 2: Stable agreements under pure multilateralism –

A comparison of Figures 1 and 2 delivers the following result:

Proposition 4: *Suppose free trade is stable when $t < t_l$ and $e^{bc} < x < x_t$. Then, (i) the freedom to pursue bilateral agreements is necessary for achieving global free trade whenever $e^{bc} < x < e^{\phi}$ and (ii) a purely multilateral approach yields the status quo instead of a bilateral trade agreement between the two larger countries whenever $x < e^{bc}$.*

Part (i) of proposition 4 hinges on the insight that when free trade is infeasible due to the reluctance of the small country (which happens when $x < e^{\phi}$), the fact that the two large countries can form a bilateral FTA can induce the small country to participate in global free trade since its welfare as non-member facing $\langle\{bc\}\rangle$ is lower than that under free trade. This result highlights the intuition that under a purely multilateral approach, a country that is reluctant to liberalize can effectively stall global trade liberalization and the freedom to pursue bilateral agreements effectively removes any single country's ability to *veto* trade liberalization between *other* countries.

When $x < e^{bc}$ global free trade is not a stable equilibrium regardless of whether countries follow pure multilateralism or not because country a is better off as an outsider facing the FTA $\langle\{bc\}\rangle$ than it is under $\langle\{F\}\rangle$. Under such a situation, bilateralism is beneficial for another reason: it leads to welfare-improving trade liberalization in the form of $\langle\{bc\}\rangle$ whereas no

agreement $\langle\{\Phi\}\rangle$ obtains under pure multilateralism. Figure 3 illustrates the beneficial effects of bilateralism.

– Figure 3: Beneficial effects of bilateralism –

We now comment on the case where $\langle\{bc\}\rangle$ is stable when $t < t_l$ and $e^{bc} < x < x_t$. Then, pure multilateralism has the following effects: (i) it undermines global free trade by yielding the status quo when either (a) $x_t < x < e^{\phi}$ or (b) $e^{bc} < x < e^{\phi}$ and $t > t_l$; (ii) it prevents an FTA between the two big countries when either (a) $x < e^{bc}$ or (b) $x < e^{\phi}$, $x < x_t$ and $t < t_l$; and (iii) it facilitates global trade liberalization by yielding $\langle\{F\}\rangle$ instead of $\langle\{bc\}\rangle$ whenever $e^{\phi} < x < x_t$ and $t < t_l$. We illustrate these results in Figure 4.

– Figure 4: Mixed effects of bilateralism –

It is worth emphasizing that pure multilateralism can act as a force in favor of global free trade by replacing the bilateral FTA between two large countries by free trade whenever $e^{\phi} < x < x_t$ and $t < t_l$. When such is the case, pure multilateralism benefits all countries. The logic for this is as follows. In the absence of pure multilateralism, while all countries are willing to deviate from $\langle\{bc\}\rangle$ to $\langle\{F\}\rangle$, this deviation is not self-enforcing because countries a and b further deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. By ruling out such a deviation, pure multilateralism facilitates the obtainment of global free trade and can prevent two large countries from signing an agreement that excludes the small country. The following proposition summarizes the distributional implications of pure multilateralism:

Proposition 5: *When pure multilateralism facilitates global free trade, it benefits all countries and when it undermines global free trade, it hurts the large countries whereas it benefits the small country. Furthermore, when global free trade is infeasible (regardless of whether countries follow strict multilateralism or not), pure multilateralism hurts the large countries whereas it benefits the small country.*

The first part of the proposition is clear given the above discussion. For the second part, note that the small country always prefers no agreement to

a bilateral agreement between the other two larger countries:

$$w_a(\Phi) - w_a(bc) = \frac{t(2x - e - 3\beta t)}{9} \geq 0 \text{ since } x > x^{cr} \tag{30}$$

Thus, when pure multilateralism preserves the status quo by ruling out such a bilateral agreement, it necessarily makes the small country better off. Recall that pure multilateralism undermines global free trade precisely when $e^{bc} < x < e^{\phi} \Leftrightarrow w_a(bc) < w_a(F) < w_a(\Phi)$. Finally, when global free trade is infeasible under both bilateralism and pure multilateralism, $\langle\{bc\}\rangle$ is the stable agreement under bilateralism whereas no agreement $\langle\{\Phi\}\rangle$ obtains under pure multilateralism. Under such a case, larger countries b and c always prefer the ability to form a bilateral FTA $\langle\{bc\}\rangle$ to no agreement $\langle\{\Phi\}\rangle$ while the opposite is true for the excluded small country (immediate from (30)):

$$w_i(bc) \geq w_i(\Phi) = \frac{t(2x - e - 3\beta t)}{18} \geq 0 \text{ since } x > x^{cr},\ i = b, c \tag{31}$$

5.2 One large and two small countries

In this section we analyze the case where countries a and b have smaller endowments than country c:

Assumption 3: $x = y = e < z$.

As before, we exclude prohibitive tariff levels by assuming the following condition: $z < z^{cr} \equiv 2e - 4\beta t$.

5.2.1 Stable agreements

We first derive conditions under which free trade is a stable equilibria. Similar to the previous case, we consider the perspective of the large country first. As might be expected, the large country is better off under free trade relative to no agreement:

$$w_c(F) \geq w_c(\Phi) \tag{32}$$

Moreover, country c has no incentive to unilaterally deviate from free trade $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$ or $\langle\{ah\}\rangle$ ($\langle\{bh\}\rangle$):

$$w_c(F) \geq w_c(ab) \tag{33}$$

and

$$w_c(F) \geq w_c(ah) = w_c(bh) \tag{34}$$

Finally, country c has an incentive to jointly deviate with one of the smaller countries (a or b) from $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$ ($\langle\{bc\}\rangle$) if the tariff level t is sufficiently small:

$$w_c(F) \leq w_c(ac) = w_c(bc) \text{ iff } t \leq \frac{e}{9\beta t} \tag{35}$$

and a small country also has the same incentive when the large country's endowment is not too big:

$$w_a(F) = w_b(F) \leq w_a(ac) = w_b(bc) \text{ iff } z \leq \frac{4e}{3} - 3\beta t \tag{36}$$

Note that if $z \leq \frac{4e}{3} - 3\beta t$ holds, then $t \leq \frac{e}{9\beta t}$ obtains (since $z > e$). Thus, in the joint deviation from $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$ ($\langle\{bc\}\rangle$) the incentive of the small member country binds. However this deviation is not self-enforcing since the large country has an incentive to deviate further from $\langle\{ac\}\rangle$ ($\langle\{bc\}\rangle$) to $\langle\{ch\}\rangle$:

$$w_c(ch) \geq w_c(ac) = w_c(bc) \tag{37}$$

Now consider the perspective of the two small countries (i.e. a and b). First note that there exists a critical threshold endowment level e^{ϕ} such that if the large country's endowment exceeds e^{ϕ} then each small country actually prefers no agreement to free trade:

$$w_i(\Phi) \geq w_i(F) \text{ iff } z \geq e^{\phi} \equiv e + 3\beta t,\ i = a, b \tag{38}$$

As a result, the small countries have an incentive to jointly deviate from free trade to no agreement if the large country's endowment is sufficiently big (i.e. when $z > e^{\phi}$). But is this deviation self-enforcing? It turns out that if the endowment of the large country falls below a certain threshold, one of of the deviating small countries (say a), has an incentive to further deviate from $\langle\{\Phi\}\rangle$ to $\langle\{ac\}\rangle$:

$$w_a(ac) \geq w_a(\Phi) \text{ iff } z \leq e^{\phi - ac} \equiv \frac{6e - 3\beta t}{5} \tag{39}$$

Therefore, if $z \leq e^{\phi - ac}$ holds, then the initial joint deviation of small countries from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$ is not self-enforcing.

Now consider the small countries' incentives to jointly deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. It turns out that if country c is sufficiently large, the small countries indeed have an incentive to jointly deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$:

$$w_a(F) - w_a(ab) = w_b(F) - w_b(ab) \leq 0 \text{ iff } z \geq e^{bc} \equiv \frac{3e + 9\beta t}{4} \tag{40}$$

However, unless country c is sufficiently large, one of the smaller countries' (say a) has an incentive to further deviate from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$:

$$w_a(ah) \geq w_a(ab) \text{ iff } z \leq e^{ah-ab} \equiv \frac{6e + 11\beta t}{5} \tag{41}$$

Therefore, if $z \leq e^{ah-ab}$ the initial joint deviation of the two small countries from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$ is not self-enforcing. It is easy to show that $e^{ah-ab} > e^{\phi}$ and $e^{ah-ab} \geq e^{\phi - ac}$ – i.e. the joint deviation from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$ does not determine the stability of free trade.

Moreover, country a has an incentive to unilaterally deviate from $\langle\{F\}\rangle$ to $\langle\{bc\}\rangle$ if large country's endowment is sufficiently large:

$$w_a(F) \leq w_a(bc) \text{ iff } z \geq e^{bc} \equiv \frac{5e}{3} \tag{42}$$

But this deviation is not binding in determining stability of free trade since $e^{bc} > \max\{e^{\phi}, e^{\phi - ac}\}$.

Next we consider unilateral deviation of a small country (say b) from $\langle\{F\}\rangle$ to a pair of bilateral FTAs where the other small country is a hub ($\langle\{ah\}\rangle$. We have:

$$w_b(F) \leq w_b(ah) \text{ if } z \geq e^{ah} \equiv \frac{6e + 7\beta t}{5} \tag{43}$$

Finally note that neither small country has an incentive to deviate from $\langle\{F\}\rangle$ to $\langle\{ch\}\rangle$ where the large country is the hub:

$$w_a(F) \geq w_a(ch) \tag{44}$$

We can now state the main result:[12]

[12] Proofs of statements (*ii*) through (*iv*) are in the appendix.

Proposition 6: *Free trade* $\langle\{F\}\rangle$ *is the unique stable equilibrium when (a)* $z < e^{\phi-ac}$ or (b) $z < \min\{e^{\phi}, e^{ah}\}$*; (ii)* $\langle\{ah\}\rangle$ *is uniquely stable if* $e^{ah} < z < e^{ah-ab}$*; (iii)* $\langle\{ab\}\rangle$ *is uniquely stable if* $e^{ah-ab} < z$*; (iv) no agreement* $\langle\{\Phi\}\rangle$*, a bilateral FTA between a large and a small country (i.e.* $\langle\{ac\}\rangle$ or $\langle\{bc\}\rangle$*) and the pair of bilateral FTAs with the large country as hub* $\langle\{ch\}\rangle$ *are never stable.*

– Figure 5: Stable agreements under case 2 –

The above proposition and figure 5 show that the multiplicity problem that existed for the case of two large and one small country no longer arises. However, the set of stable equilibria is now empty for an intermediate range of asymmetry and tariffs (i.e. when $z > e^{\phi-ac}$ and $e^{\phi} < z < e^{ah}$ hold simultaneously). To understand why this happens, first note that over this range the joint deviation of countries a and b from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$ is self-enforcing. Second, countries a and c have an incentive to deviate from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$ and this deviation is also self-enforcing over the given range. So, $\langle\{ab\}\rangle$ is not stable as well. Finally, consider the joint deviation of countries b and c from $\langle\{ah\}\rangle$ to $\langle\{F\}\rangle$. Over the given range, this joint deviation is self-enforcing so that $\langle\{ah\}\rangle$ (or $\langle\{bh\}\rangle$) is not stable either. As a result, when $z > e^{\phi-ac}$ and $e^{\phi} < z < e^{ah}$, there exists no stable equilibrium.

5.2.2 The role of bilateralism

What light does the above analysis shed on the pros and cons of a purely multilateral approach to trade liberalization when two countries are small relative to the third? First, note that under a purely multilateral, global free trade is a stable equilibrium only when $z < e^{\phi}$ (with status quo prevailing over the rest of the parameter space – see Figure 6).

– Figure 6: Stable multilateral agreements –

As a result, if $e^{\phi} < z < e^{\phi-ac}$, pure multilateralism ends up undermining global free trade. Intuitively, when $e^{\phi} < z < e^{\phi-ac}$ the two small countries (a and b) benefit from jointly deviating from free trade to no agreement.

However, when bilateral agreements are possible this deviation is *not* self-enforcing since each small country has an incentive to further deviate from no agreement to a bilateral FTA with the large country. By contrast, the initial joint deviation of the two small countries from free trade to no agreement is self-enforcing under a purely multilateral approach since the only alternative to free trade is no agreement.

Next note that if $e^{ah} < z < e^{\phi}$, a purely multilateral approach facilitates the obtainment of global free trade because a small country benefits from a unilateral deviation from free trade to a hub and spoke arrangement with it as a spoke and the other small country as a hub whereas it has no incentive to deviate from free trade to no agreement. Therefore, a multilateral approach proves conducive to making global free trade stable when $e^{ah} < z < e^{\phi}$.

Finally, when $z > \max\{e^{ah}, e^{\phi}\}$, free trade is not feasible regardless of whether countries follow pure multilateralism or not. Under such a situation, as in the previous case, the option to pursue bilateral agreements can yield welfare-improving trade liberalization that is foregone under the purely multilateral approach. More specifically, a purely multilateral approach prevents the emergence of (i) the hub and spoke arrangement $\langle\{ah\}\rangle$ (or $\langle\{ah\}\rangle$) when the size asymmetry across countries is of intermediate magnitude; and (ii) the bilateral FTA $\langle\{ab\}\rangle$ when the size asymmetry across countries is sufficiently pronounced.

We now comment on the distributional implications of pure multilateralism when two countries are small relative to the third. First, as in the previous case - one small and two large countries, when pure multilateralism undermines global free trade, it hurts the large country whereas it benefits small countries. Over the given endowment and tariff range, small countries (a and b) prefer no agreement to free trade and are able to secure it via a simple unilateral or joint veto to free trade under a purely multilateral approach while this veto is not self-enforcing under bilateralism. On the other hand, when pure multilateralism facilitates global free trade, the implication of pure multilateralism is reversed: it benefits the large country while it hurts the small countries. Note that both small spoke (country b) and small hub (country a) prefer $\langle\{ah\}\rangle$ to $\langle\{F\}\rangle$ while large spoke (c) is worse off under $\langle\{ah\}\rangle$ relative to free trade.

6 Concluding remarks

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 generally tended to treat bilateral trade agreements as exogenous. This is clearly unsatisfactory since we need to understand why countries enter into bilateral agreements when multilateral trade liberalization is an option. By contrast, we present a model in which 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.

We show that when only one country favors no agreement to free trade, a strictly multilateral approach is problematic because it allows such a country to effectively stall multilateral liberalization by voting against it. Under such circumstances, bilateralism can actually provide an impetus to multilateral trade liberalization – when bilateral agreements are possible, a country that chooses not to liberalize can be worse off relative to the status quo if other countries choose to liberalize trade between themselves. On the other hand, when two countries are better off under free trade relative to no agreement but have an incentive to exclude the third country, a strictly multilateral approach is relatively more conducive to the cause of global free trade. However, in our model, such an effect obtains over a rather small parameter space: more often than not, the pursuit of bilateral agreements leads to trade liberalization that is foregone under a purely multilateral approach while also reducing the likelihood of ending up with no liberalization whatsoever.

Our 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 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 is 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.

7 Appendix

7.1 Proof of lemma 2

Recall that the condition $x > x^{cr} = \frac{e}{2} + 2\beta t$ excludes prohibitive tariff levels. There are three possible unilateral deviations of large countries to consider. Since countries b and c are symmetric, it is sufficient to consider only country b's deviations:

- UF1: From $\langle\{F\}\rangle$ to $\langle\{ah\}\rangle$.
- UF2: From $\langle\{F\}\rangle$ to $\langle\{ch\}\rangle$.
- UF3: From $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$.

Note that UF1 is ruled out since

$$w_b(F) - w_b(ah) = \frac{t(2x - e + 7\beta t)}{18} \geq 0 \text{ since } x \geq x^{cr} \tag{45}$$

Similarly, UF2 is ruled out:

$$w_b(F) - w_b(ch) = \frac{t(6e - 5x + 7\beta t)}{18} \geq 0 \text{ since } e \geq x \tag{46}$$

Finally, UF3 does not occur since:

$$w_b(F) - w_b(ac) = \frac{t(5e - 3x)}{18} \geq 0 \text{ since } e \geq x \tag{47}$$

Now consider the following possible joint deviations:

- JF1: Deviation of b and c from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$.
- JF2: Deviation of b and c from $\langle\{F\}\rangle$ to $\langle\{ah\}\rangle$.
- JF3: Deviation of a and b from $\langle\{F\}\rangle$ to $\langle\{ch\}\rangle$.
- JF4: Deviation of a and b from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$.
- JF5: Deviation of b and c from $\langle\{F\}\rangle$ to $\langle\{bc\}\rangle$.

Since large countries prefer $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$, deviation JF1 is ruled out. Similarly, due to inequalities (45) and (46), JF2 and JF3 are ruled out since large countries always prefers $\langle\{F\}\rangle$ to being a spoke under a pair of bilateral FTAs (as b in $\langle\{ah\}\rangle$ or $\langle\{ch\}\rangle$). Now consider JF4. We have:

$$w_a(F) - w_a(ab) = \frac{t(-e + 9\beta t)}{18} \leq 0 \text{ iff } t < t_l \tag{48}$$

and

$$w_b(F) - w_b(ab) = \frac{t(3x - 4e + 9\beta t)}{18} \leq 0 \text{ iff } t < \frac{4e - 3x}{9\beta} \tag{49}$$

As a result, JF4 occurs if $t < t_l$. But this deviation is not self enforcing? To see this, consider the further deviation of country b from $\langle\{ab\}\rangle$ to $\langle\{bh\}\rangle$:

$$w_b(bh) - w_b(ab) = \frac{t(2x - e + 11\beta t)}{18} \geq 0 \text{ since } x \geq x^{cr} \tag{50}$$

Since country b benefits from further deviating from $\langle\{ab\}\rangle$ to $\langle\{bh\}\rangle$, the initial deviation JF4 is not self-enforcing.

Finally, consider JF5. This deviation occurs when $x > \frac{3e+9\beta t}{4}$ However, even if this deviation were to occur, country b has an incentive to further deviate from $\langle\{bc\}\rangle$ to $\langle\{bh\}\rangle$:

$$w_b(bh) - w_b(bc) = \frac{t(6e - 5x + 11\beta t)}{18} \geq 0 \text{ since } e \geq x \tag{51}$$

Therefore, the initial deviation JF5 is not self-enforcing.

7.2 Proof of proposition 3

We know that $\langle\{F\}\rangle$ is stable iff $x \geq e^{bc}$. In order to obtain entire set of stable equilibria, consider $\langle\{\Phi\}\rangle$ first. The large countries (b and c) have an incentive to form $\langle\{bc\}\rangle$ and this deviation is self-enforcing since $\langle\{bc\}\rangle$ is a Nash equilibrium:

$$w_b(bc) - w_b(\Phi) = \frac{t(2x - e - 3\beta t)}{18} \geq 0 \text{ since } x > x^{cr} = \frac{e}{2} + 2\beta t \qquad (52)$$

Therefore, $\langle\{\Phi\}\rangle$ is not stable.

Now consider $\langle\{ab\}\rangle$.[13] First, note that $\langle\{ab\}\rangle$ is not even a Nash equilibrium when country a is sufficiently small relative to its partner:

$$w_a(ab) - w_a(\Phi) = \frac{t(4x - 3e - 3\beta t)}{18} \leq 0 \text{ when } x \leq \frac{3e + 3\beta t}{4} \qquad (53)$$

Second, we know from (47) that the country c always prefers $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. Moreover, countries a and b have an incentive to jointly deviate from $\langle\{ab\}\rangle$ to $\langle\{F\}\rangle$ when country a is sufficiently large:

$$w_b(F) - w_b(ab) = \frac{t(3x - 4e + 9\beta t)}{18} \geq 0 \text{ when } x \geq \frac{4e}{3} - 3\beta t \qquad (54)$$

Given Lemma 2 and inequality in (29) above, this deviation is self-enforcing. Finally, countries a and c have an incentive to jointly deviate from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$ when a is neither sufficiently large nor small:

$$w_a(ah) - w_a(ab) = \frac{t(4x - 3e + 11\beta t)}{18} \geq 0 \text{ when } x \geq \frac{3e - 11\beta t}{4} \qquad (55)$$

and

$$w_c(ah) - w_c(ab) = \frac{t(6e - 5x - 7\beta t)}{18} \geq 0 \text{ when } x \leq \frac{6e - 7\beta t}{5} \qquad (56)$$

The only meaningful further deviation to consider is that of country a from $\langle\{ah\}\rangle$ to $\langle\{\Phi\}\rangle$. We have:

$$w_a(ah) - w_a(\Phi) = \frac{t(4x - 3e + 4\beta t)}{9} \leq 0 \text{ when } x \leq \frac{3e}{4} - \beta t \qquad (57)$$

[13] Due to symmetry, the same analysis applies to $\langle\{ac\}\rangle$.

It follows that the initial joint deviation from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$ is self-enforcing when $\frac{3e}{4} - \beta t \leq x \leq \frac{6e-7\beta t}{5}$ holds. As a result, combining this condition with (53) and (54), $\langle\{ab\}\rangle$ is never stable.

Now, consider $\langle\{bc\}\rangle$. We know from (52) that $\langle\{bc\}\rangle$ is a Nash equilibrium. Moreover, we know from (50) and (45) that the joint deviations from $\langle\{bc\}\rangle$ to $\langle\{ab\}\rangle$ (or $\langle\{ac\}\rangle$) and $\langle\{ah\}\rangle$ are not self enforcing. Next, consider the deviation of a and b from $\langle\{bc\}\rangle$ to $\langle\{bh\}\rangle$. It is immediate from (51) that country b prefers $\langle\{bh\}\rangle$ to $\langle\{bc\}\rangle$ which implies that country a's choice is binding with respect to this joint deviation:

$$w_a(bh) - w_a(bc) = \frac{t(4x - 3e - 7\beta t)}{18} \geq 0 \text{ when } x \geq \frac{3e + 7\beta t}{4} \tag{58}$$

Thus, this joint deviation is self-enforcing since neither country a nor b has an incentive to further deviate from $\langle\{bh\}\rangle$.

Now consider the joint deviation of all countries from $\langle\{bc\}\rangle$ to $\langle\{F\}\rangle$. We know from (29) that this joint deviation occurs only if $x > e^{bc}$. But this joint deviation is not self-enforcing when $t < t_l$ since countries a and b have an incentive to further deviate from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$. Thus, we have shown $\langle\{bc\}\rangle$ is stable if (i) $x \leq e^{bc}$ or (ii) $t < t_l$ when $x_t > x > e^{bc}$.

Finally, consider the hub and spoke trade agreements. First note that countries b and c always have an incentive to jointly deviate from $\langle\{ah\}\rangle$ to $\langle\{F\}\rangle$ and this deviation is self enforcing since neither b nor c has an incentive to unilaterally deviate from free trade $\langle\{F\}\rangle$ (see Lemma 2). As a result, $\langle\{ah\}\rangle$ is never stable. Second, country a has an incentive to unilaterally deviate from $\langle\{bh\}\rangle$ to $\langle\{bc\}\rangle$ if $x < x_t$. On the other hand, when $x \geq x_t$ holds, countries a and c have an incentive to jointly deviate from $\langle\{bh\}\rangle$ to $\langle\{F\}\rangle$ and this joint deviation is self-enforcing. As a result, $\langle\{bh\}\rangle$ is never stable as well.

7.3 Proof of proposition 5

When $e^{\phi} < x < x_t$ and $t < t_l$ holds, pure multilateralism facilitates global free trade by replacing the bilateral FTA between two large countries $\langle\{bc\}\rangle$ by free trade $\langle\{F\}\rangle$. Over this range all countries prefer $\langle\{F\}\rangle$ to $\langle\{bc\}\rangle$. First consider the small country:

$$w_a(F) - w_a(bc) = \frac{t(4x - 3e)}{9} \geq 0 \text{ iff } x \geq e^{bc} \equiv \frac{3e}{4} \tag{59}$$

Since $e^{\phi} \geq e^{bc}$, over the given range ($e^{\phi} < x < x_t$) the inequality in (59) always holds. Now consider the large countries' perspective:

$$w_i(F) - w_i(bc) = \frac{t(3e - 4x + 9\beta t)}{18} \geq 0 \text{ iff } x \leq e_L^{bc} = \frac{3e + 9\beta t}{4},\ i = b, c \tag{60}$$

Note that $e_L^{bc} \geq x_t \equiv \frac{3e+7\beta t}{4}$ always holds. Therefore, it is immediate to argue that over the range $e^{\phi} < x < x_t$, large countries always prefer $\langle\{F\}\rangle$ to $\langle\{bc\}\rangle$. The proof of the second part of the proposition is immediate from the text. For the final part of the proposition, note that when global free trade is infeasible (regardless of whether countries follow strict multilateralism or not), larger countries b and c always prefer bilateralism whereas the small country prefers pure multilateralism:

$$w_a(\Phi) - w_a(bc) = \frac{t(2x - e - 3\beta t)}{9} \geq 0 \text{ since } x > x^{cr}$$

$$w_i(bc) \geq w_i(\Phi) = \frac{t(2x - e - 3\beta t)}{18} \geq 0 \text{ since } x > x^{cr},\ i = b, c \tag{61}$$

7.4 Proof of proposition 6

The range over which $\langle\{F\}\rangle$ is stable is immediate from the text. In order to obtain entire set of stable equilibria, consider $\langle\{\Phi\}\rangle$ first. The small countries (a and b) have an incentive to form $\langle\{ab\}\rangle$ and this deviation is self-enforcing since $\langle\{ab\}\rangle$ is a Nash equilibrium:

$$w_i(ij) - w_i(\Phi) = \frac{t(2z - e - 3\beta t)}{18} \geq 0,\ i, j = a, b. \tag{62}$$

Therefore, $\langle\{\Phi\}\rangle$ is not stable.

Bilateral FTA $\langle\{ac\}\rangle$ (or $\langle\{bc\}\rangle$) is Nash equilibrium iff country c is not sufficiently large:

$$w_i(ic) - w_i(\Phi) = \frac{t(6e - 5z - 3\beta t)}{18} \leq 0 \text{ iff } z \geq e^{\phi - ic} \equiv \frac{6e - 3\beta t}{5}\ i = a, b. \tag{63}$$

Note also that small countries (a and b) have an incentive to jointly deviate from $\langle\{ac\}\rangle$ to $\langle\{ah\}\rangle$ (or $\langle\{bc\}\rangle$ to $\langle\{bh\}\rangle$) unless country c is sufficiently small and tariff level is sufficiently high and this deviation is self-enforcing:

$$w_i(ih) - w_i(ic) = \frac{t(2z - e + 11\beta t)}{18} \geq 0 \text{ since } z > e,\ i = a, b. \tag{64}$$

$$w_j(ih) - w_j(ic) = \frac{t(2z - e - 7\beta t)}{18} \geq 0 \text{ iff } z \geq e^{ih-ic} \equiv \frac{e + 7\beta t}{2},\ i, j = a, b. \tag{65}$$

The inequalities in (63) and (65) cover entire parameter space. Thus, the bilateral FTA between large and small countries ($\langle\{ac\}\rangle$ or $\langle\{bc\}\rangle$) is never stable.

Now consider the pair of bilateral FTAs where the large country is hub $\langle\{ch\}\rangle$. A small spoke (say $i = a, b$) has an incentive to deviate unilaterally from $\langle\{ch\}\rangle$ to $\langle\{jc\}\rangle$ when country c is sufficiently large:

$$w_i(ch) - w_i(jc) = \frac{t(6e - 5z - 7\beta t)}{18} < 0 \text{ iff } z > e^{ch-jc} \equiv \frac{6e - 7\beta t}{5},\ i, j = a, b. \tag{66}$$

Note that when $z \leq e^{ch-jc}$ holds, two small spokes deviate from $\langle\{ch\}\rangle$ to $\langle\{F\}\rangle$ and this deviation is self-enforcing since $e^{ch-jc} < e^{ah}$. As a result, $\langle\{ch\}\rangle$ is never stable.

Next, we consider pair of bilateral FTAs where a small country is hub (say $\langle\{ah\}\rangle$). When country c is sufficiently large, small hub (country a) unilaterally deviates to a bilateral agreement $\langle\{ab\}\rangle$ where large country is excluded:

$$w_a(ah) - w_a(ab) = \frac{t(6e - 5z + 11\beta t)}{18} < 0 \text{ iff } z > e^{ah-ab} \equiv \frac{6e + 11\beta t}{5} \tag{67}$$

Moreover, when $z \leq e^{ah}$ holds, spoke countries (b and c) jointly deviate to free trade $\langle\{F\}\rangle$ and it is a self-enforcing deviation. These two deviations determine the stability region: $\langle\{ih\}\rangle$ ($i = a, b$) is stable iff $e^{ah} \leq z \leq e^{ah-ab}$.

Finally consider a bilateral agreement of two small countries $\langle\{ab\}\rangle$. First note from (67) that when $z < e^{ah-ab}$ holds, country a has an incentive to deviate from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$. Similarly, country c has an incentive to

deviate from $\langle\{ab\}\rangle$ to $\langle\{ah\}\rangle$ unless countries are relatively symmetric and tariff is sufficiently high:

$$w_c(ah) - w_c(ab) = \frac{t(4z - 3e - 7\beta t)}{18} > 0 \text{ iff } z > e_c^{ah-ab} \equiv \frac{3e + 7\beta t}{4} \quad (68)$$

Thus, joint deviation happens when $e_c^{ah-ab} < z < e^{ah-ab}$ and it is a self-enforcing deviation. Note that when $z \leq e_c^{ah-ab}$, free trade is immune to any unilateral or joint deviations (strong Nash equilibrium). As a result, $\langle\{ab\}\rangle$ is stable iff $z \geq e^{ah-ab}$ holds.

7.5 Supporting calculations

7.6 Symmetry

We calculate:

$w_i(ij) - w_i(\Phi) = \frac{t(e-3\beta t)}{18} \geq 0$ since $t < \overline{t} = \frac{e}{4\beta}$.
$w_i(ih) - w_i(F) = \frac{t(e+\beta t)}{9} \geq 0$.
$w_j(ih) - w_j(F) = -\frac{t(e+7\beta t)}{18} \leq 0$.
$w_i(ih) - w_i(\Phi) = \frac{t(e+4\beta t)}{9} \geq 0$.
$w_j(ih) - w_j(ik) = \frac{t(e-7\beta t)}{18} \geq 0$ iff $t \leq t_h \equiv \frac{e}{7\beta}$.
$w_i(F) - w_i(jk) = \frac{te}{9\beta} \geq 0$.
$w_i(F) - w_i(ij) = \frac{t(-e+9\beta t)}{18} \geq 0$ iff $t \geq t_l \equiv \frac{e}{9\beta}$.
$w_i(ih) - w_i(ij) = \frac{t(e+11\beta t)}{18} \geq 0$.

7.7 One small and two large countries

We have:

$w_a(\Phi) - w_a(F) = \frac{t(2e-2x-3\beta t)}{9} \geq 0$ iff $x \leq e^{\phi} \equiv \frac{2e-3\beta t}{2}$.
$w_a(bh) - w_a(F) = \frac{t(3e-4x-7\beta t)}{18} \geq 0$ iff $x \leq e^{bh} \equiv \frac{3e-7\beta t}{4}$.
$w_a(bc) - w_a(F) = \frac{t(3e-4x)}{9} \geq 0$ iff $x \leq e^{bc} \equiv \frac{3e}{4}$.
Since $x > x^{cr}$ holds, the following always obtains: $e^{\phi} > e^{bc} > e^{bh}$

7.8 One large and two small countries

We have:

$w_c(F) - w_c(\Phi) = \frac{t(2z-2e+3\beta t)}{9} \geq 0$ since $z > e$.

$w_i(\Phi) - w_i(F) = \frac{t(z-e-3\beta t)}{9} \geq 0$ iff $z \geq e^{\phi} \equiv e + 3\beta t$, $i = a, b$.
$w_c(F) - w_c(ab) = \frac{t(4z-3e)}{9} \geq 0$ since $z > e$.
$w_c(F) - w_c(ih) = \frac{t(7\beta t+4z-3e)}{18} \geq 0$ since $z > e$, $i = a, b$.
$w_c(F) - w_c(ic) = \frac{t(-e+9\beta t)}{18} \leq 0$ iff $t \leq \frac{e}{9\beta t}$, $i = a, b$.
$w_i(F) - w_i(ic) = \frac{t(3z-4e+9\beta t)}{18} \leq 0$ iff $z \leq \frac{4e}{3} - 3\beta t$, $i = a, b$.
$w_i(ih) - w_i(ij) = \frac{t(6e-5z+11\beta t)}{18} \geq 0$ iff $z \leq e^{ih-ij} \equiv \frac{6e+11\beta t}{5}$, $i, j = a, b$.
$w_i(ih) - w_i(ic) = \frac{t(2z-e+11\beta t)}{18} \geq 0$ since $z > e$, $i = a, b$.
$w_i(ic) - w_i(\Phi) = \frac{t(6e-5z-3\beta t)}{18} \geq 0$ iff $z \leq e^{\phi-ic} \equiv \frac{6e-3\beta t}{5}$, $i = a, b$.
$w_i(F) - w_i(ij) = \frac{t(3e-4z+9\beta t)}{18} \leq 0$ iff $z \geq e^{ij} \equiv \frac{3e+9\beta t}{4}$, $i, j = a, b$.
$w_i(F) - w_i(jc) = \frac{t(5e-3z)}{18} \leq 0$ iff $z \geq e^{jc} = \frac{5e}{3}$, $i, j = a, b$.
$w_i(F) - w_i(ih) = \frac{t(3e-z+2\beta t)}{18} \geq 0$ since $z < z^{cr}$, $i = a, b$.
$w_i(F) - w_i(jh) = \frac{t(6e-5z+7\beta t)}{18} \leq 0$ if $z \geq e^{ih} = \frac{6e+7\beta t}{5}$, $i, j = a, b$.
$w_i(F) - w_i(ch) = \frac{t(2z-e+7\beta t)}{18} \geq 0$ since $z > e$, $i = a, b$.

References

[1] Aghion, Philippe, Pol Antràs, and Elhanan Helpman, 2004. "Negotiating Free Trade." NBER Working Paper 10721.

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

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

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

[6] 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.

[7] Brander, James A., and Barbara J. Spencer. "Tariff Protection and Imperfect competition." In ed. H. Kierzkowski Monopolistic Competition and International Trade, 1984, Oxford University Press, Oxford.

[8] Chang, Won and Winters, L. Alan, 2002. "How Regional Blocs Affect Excluded Countries: The Price Effects of MERCOSUR." American Economic Review 92, 889-904.

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

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

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

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

[13] 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.

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

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

[16] Ornelas, Emanuel, 2006. "Rent Destruction and the Political Viability of Free Trade Agreements." Quarterly Journal of Economics, forthcoming.

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

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

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

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

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

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

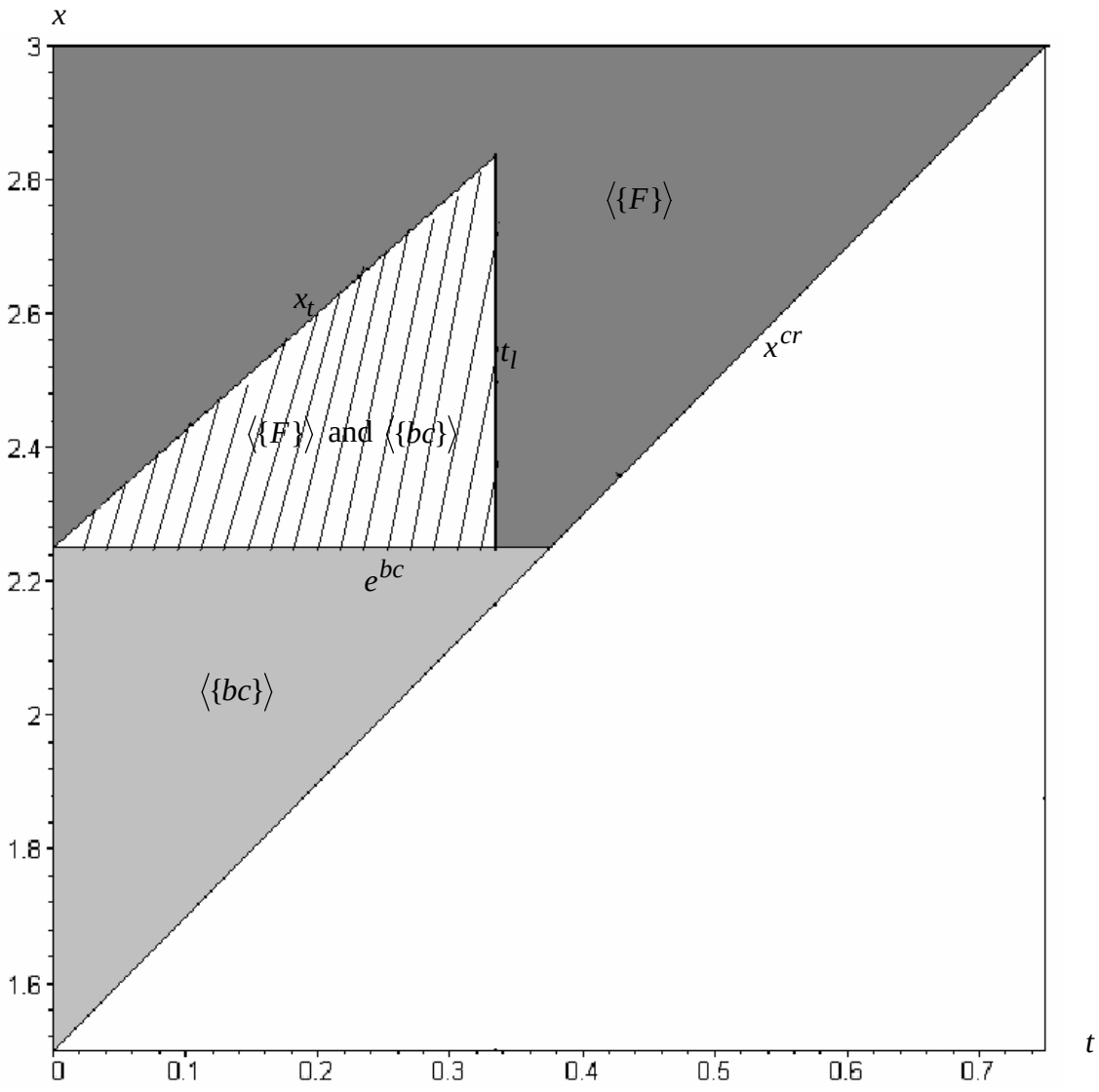

Figure 1: Stable agreements

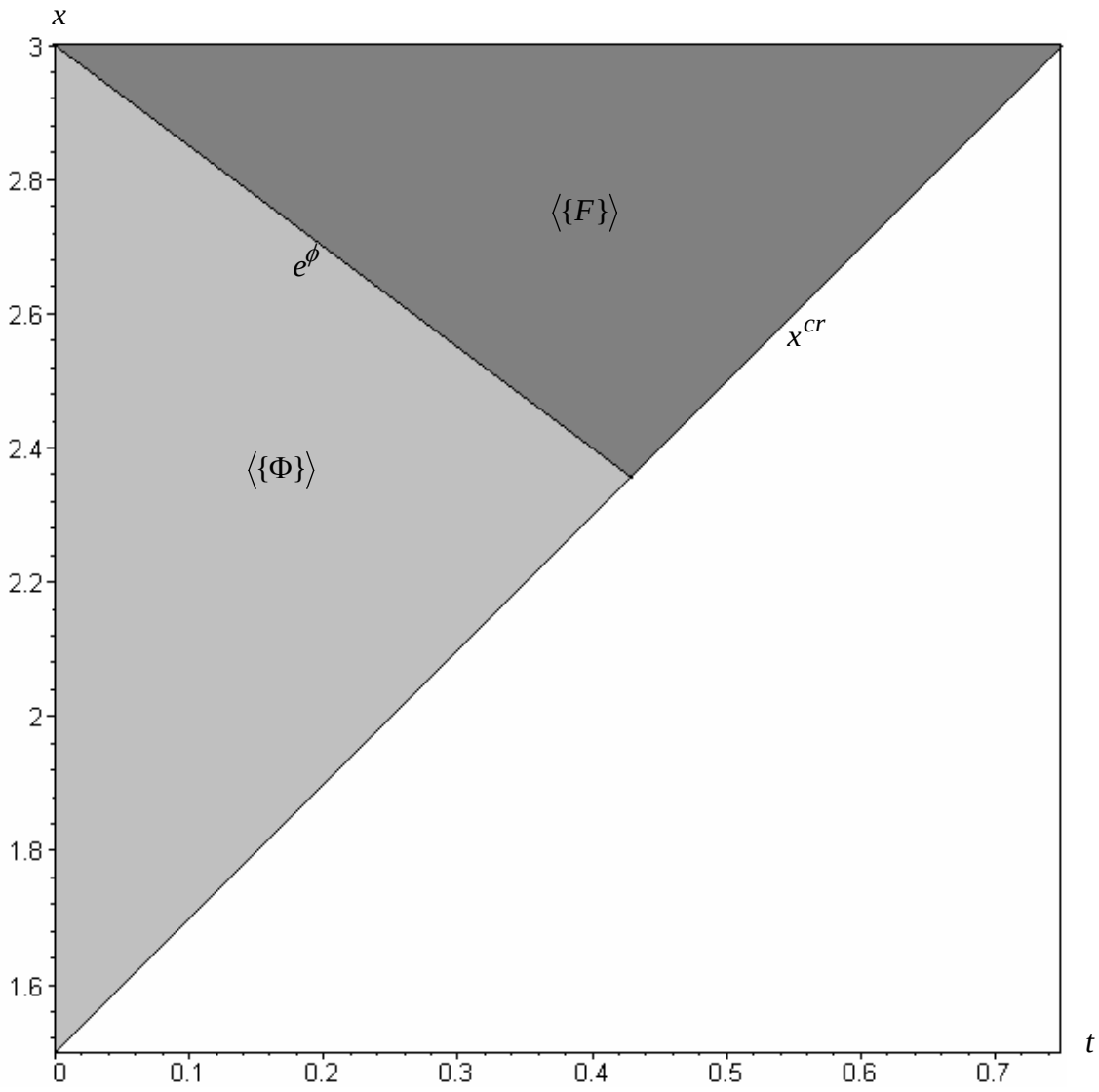

Figure 2: Stable agreements under pure multilateralism

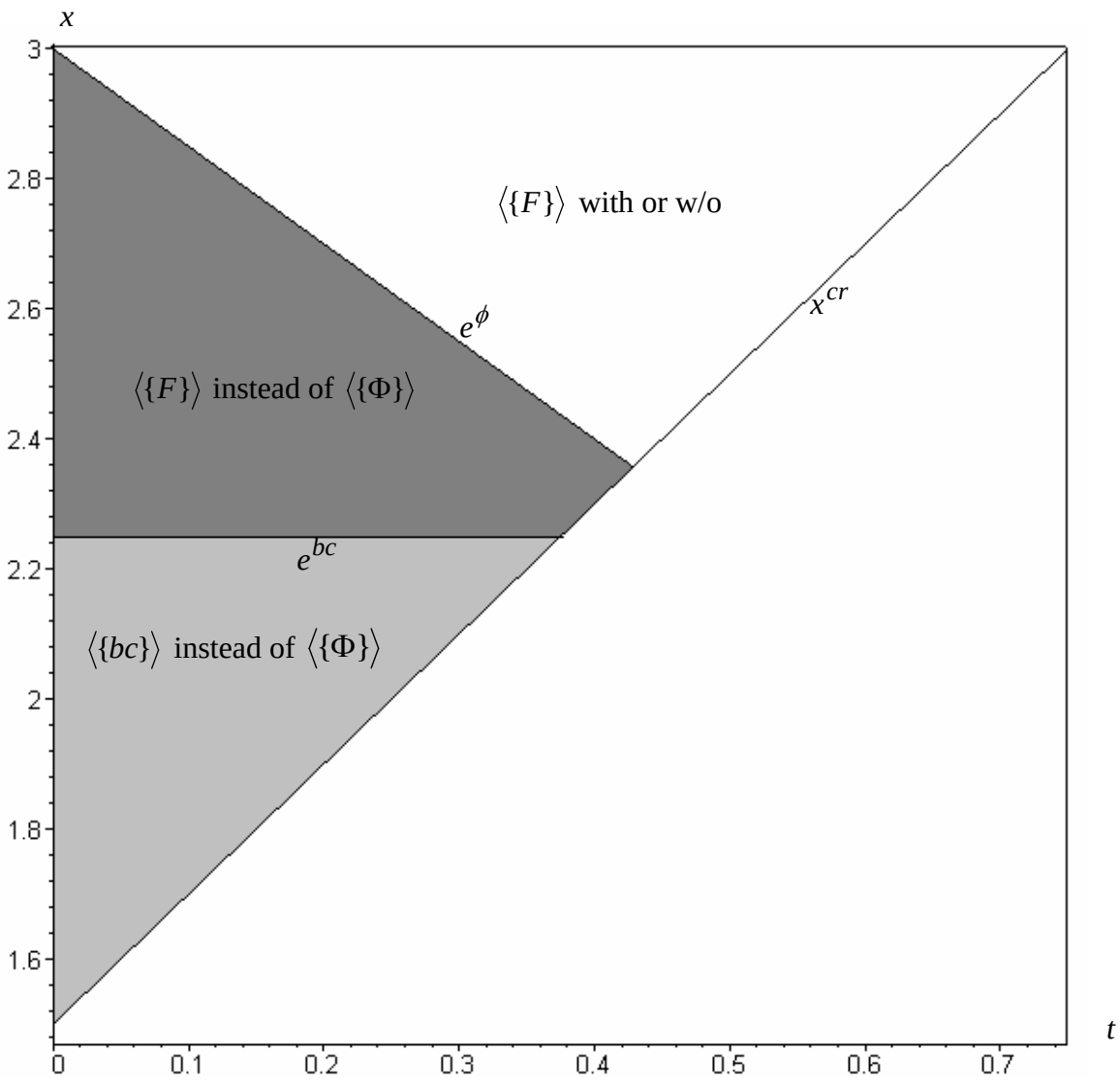

Figure 3: Beneficial bilateralism

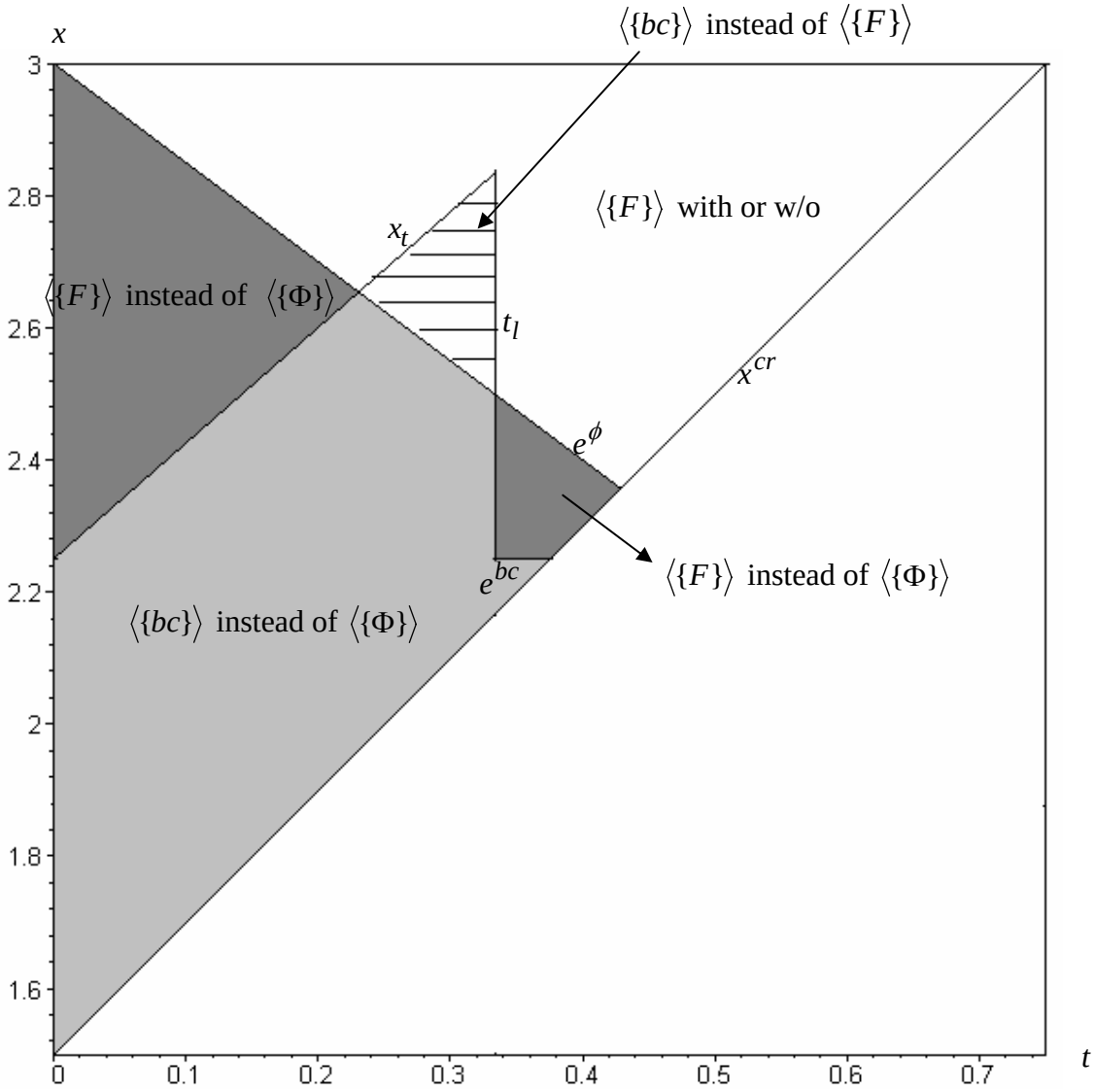

Figure 4: Mixed effects of bilateralism

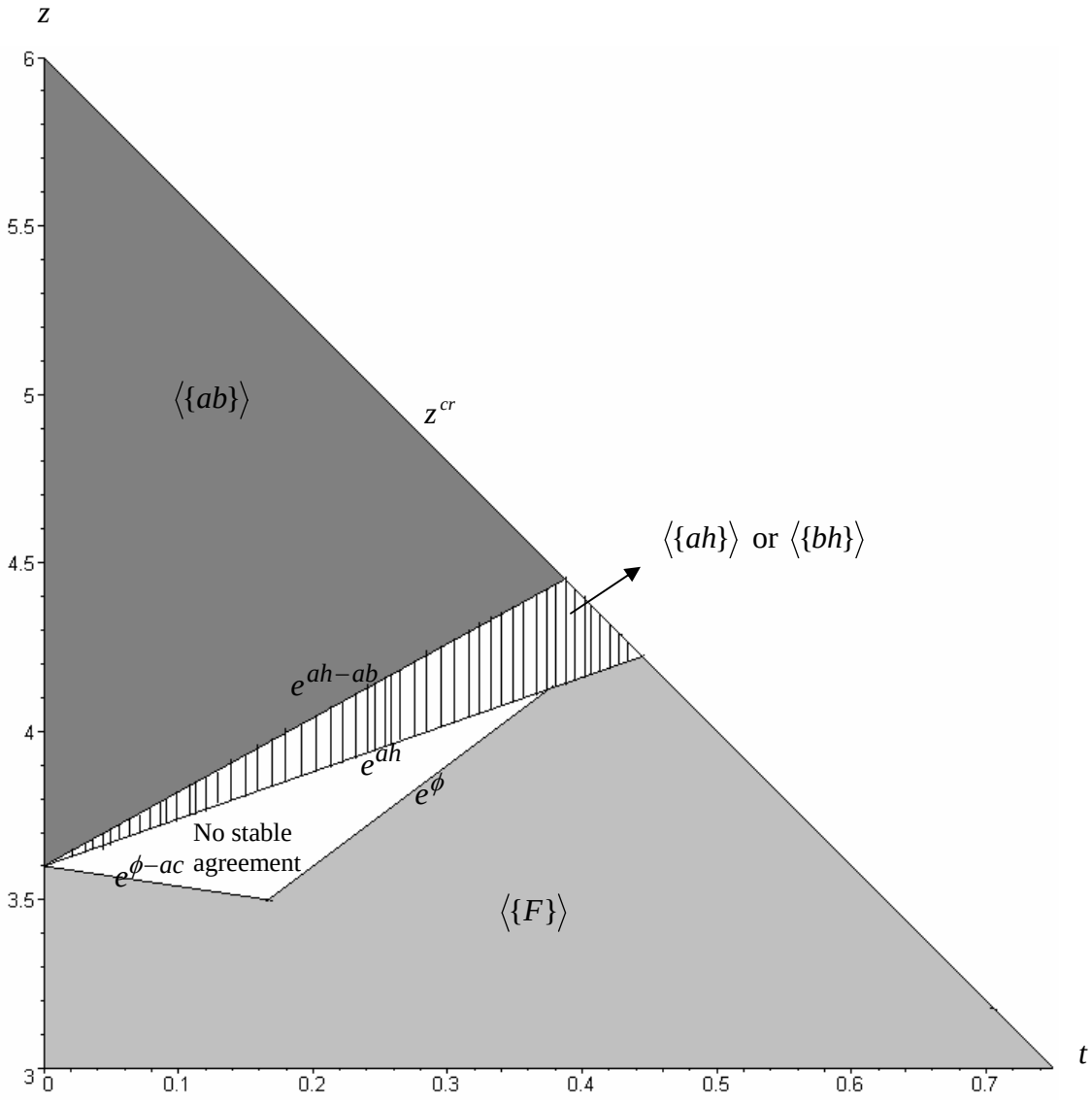

Figure 5: Stable agreements under case 2

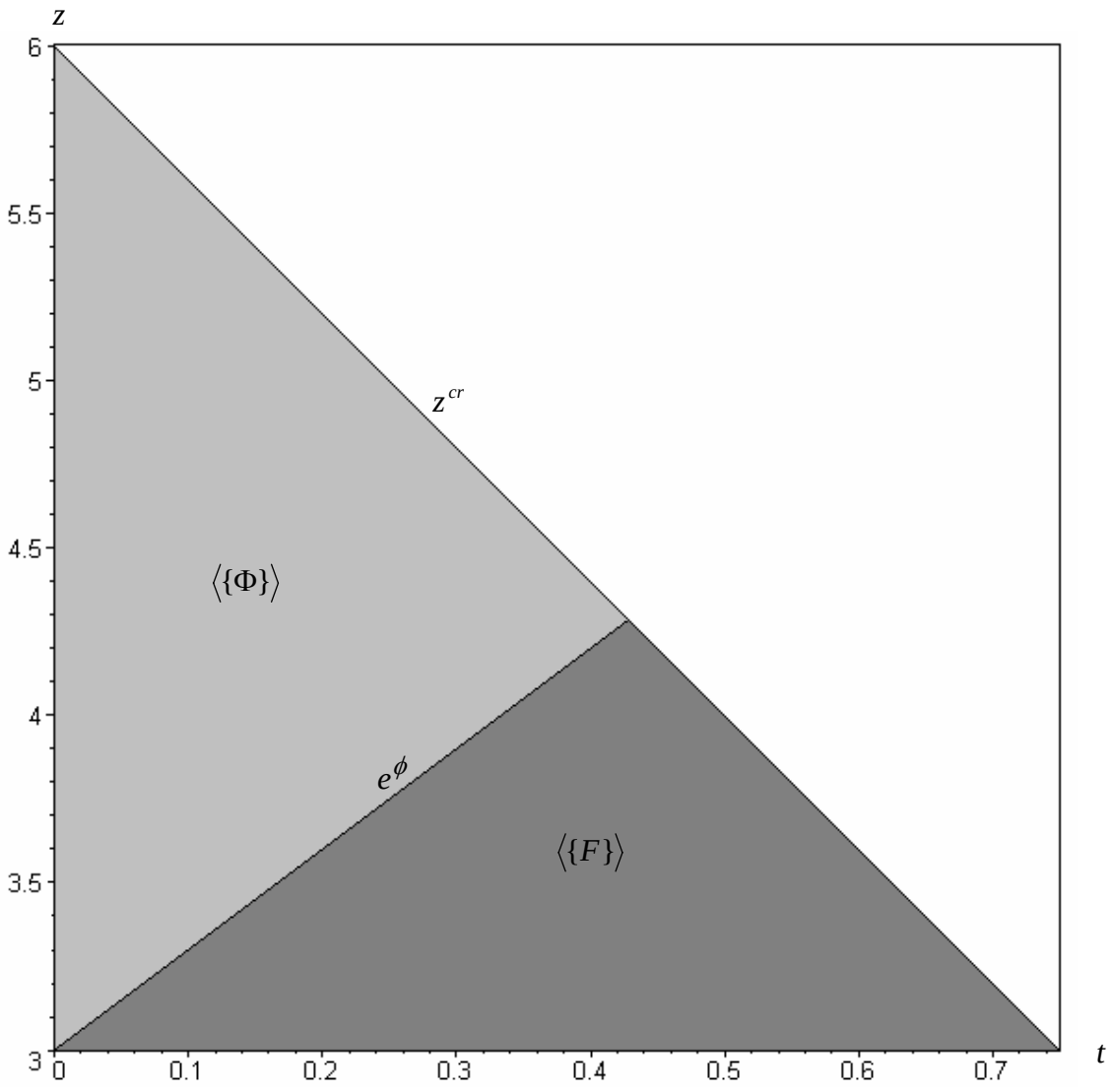

Figure 6: Stable trade agreements under pure multilateralism

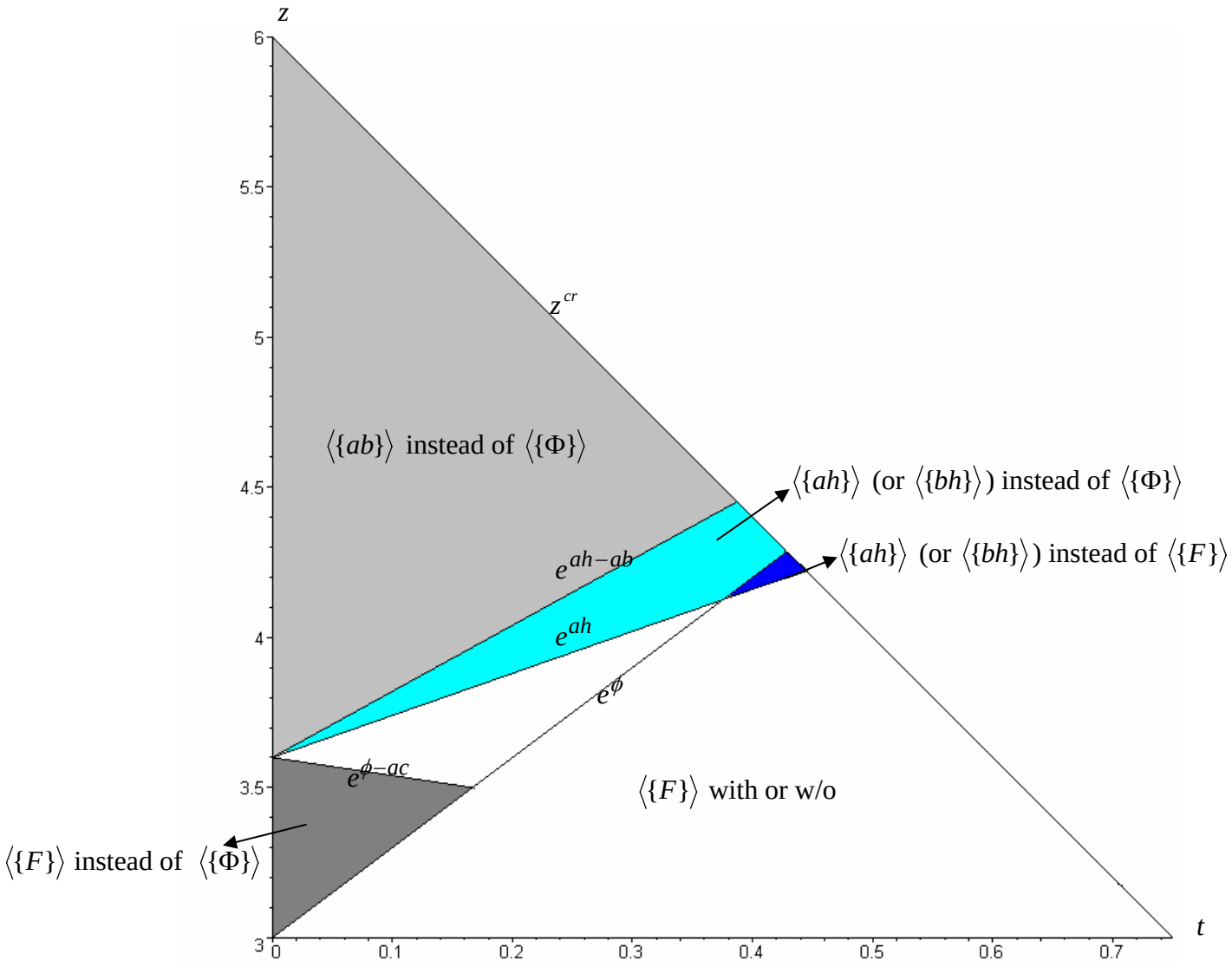

Figure 7: Effects of bilateralism under case 2