Bilateral trade agreements and the feasibility of multilateral free trade

Saggi, Kamal and Yildiz, Halis Murat

Southern Methodist University, Ryerson University

1 February 2008

Bilateral trade agreements and the feasibility of multilateral free trade

Abstract

This paper compares stable Nash equilibria of two games of trade liberalization. In the FTA game, each country can form an FTA with either one of its trade partners, or both of them, or none of them. By contrast, in the No FTA game, each country must choose either no agreement or free trade. Under symmetry, free trade is uniquely stable under the No FTA game whereas the FTA game also admits a bilateral FTA as an equilibrium. However, there exist patterns of cost asymmetry for which the freedom to pursue bilateral FTAs is necessary for achieving global free trade.

Keywords: Multilateral Trade Liberalization, Free Trade Agreements, GATT, Intraindustry Trade, Oilgopoly. JEL Classifications: F13, F12.

1 Introduction

By their very nature, preferential trade agreements (PTAs) require member countries to grant tariff reductions to each other that are typically not extended to non-members. Ever since Jacob Viner's (1950) classic analysis, the static distortions created by such preferential trade liberalization have received significant attention from economists and policy-makers alike. Furthermore, in recent years there has been widespread concern regarding the potential adverse effects of PTAs on the process of multilateral trade liberalization – the *raison d'etre* of the World Trade Organization (WTO). This concern appears to be rather well-founded: so widespread are PTAs today that Mongolia is the only country in the world that does not belong to one. As per the WTO's web-site, over 200 PTAs are officially in force today and their number is expected to reach 400 by 2010. Furthermore, such arrangements have spread rather rapidly in the last decade or so: since 1996, as many as 150 new PTAs have come into existence. Under the intricate and ever-increasing web of PTAs, the notion of most favored nation (MFN) treatment has begun to appear more of an exception rather than a core rule of the WTO.

The two most commonly occurring PTAs are free trade agreements (FTAs) and customs unions, with over 80% them being FTAs (Crawford and Fiorentino, 2005). Accordingly, this paper focuses on FTAs and asks: how does the pursuit of FTAs interact with the process of multilateral trade liberalization? Would global free trade be easier to achieve if countries were to pursue trade liberalization only multilaterally? Or as Jagdish Bhagwati (1991) put it – are FTAs building or stumbling blocs for multilateral trade liberalization? While the meaning of the phrase 'stumbling bloc' is relatively clear, what does the phrase 'building bloc' precisely mean? Does it mean that the process of bilateral trade liberalization eventually converges to multilateral free trade (as in Furusawa and Konishi (2007) and Goyal and Joshi (2006))? Or does it mean that FTAs lay the foundation for multilateral trade liberalization in the sense that the freedom to pursue FTAs is *necessary to* attain global free trade? Existing literature has often tended to take the view that FTAs are building blocs so long as their pursuit *does not prevent or eventually leads* to the obtainment of global free trade. We take this argument one step further and show that there exist circumstances where global free trade is an equilibrium *only if* countries are free to form bilateral FTAs. When such is the case, we say that FTAs act as *strong building blocs*. To our knowledge, with the exception of Aghion et. al. (2007), this potential role of FTAs has been overlooked by the voluminous existing literature on the subject. The reason for this is easy to see – as we explain below, only a model in which both bilateral and multilateral negotiations are endogenous can lead to such an insight.

The general idea behind our strong building bloc result can be stated as follows. If bilateral FTAs are prohibited by multilateral rules (or infeasible due to some other reasons) and the choice is only between multilateral free trade or no agreement (i.e. the status quo), any single country can ensure that the status quo prevails by simply opting to not practise free trade itself. However, when the rest of the world is free to form bilateral FTAs, a country that makes such a choice can find itself immiserized relative to the status quo if the other countries choose to undertake preferential trade liberalization amongst themselves.[1] Anticipating this outcome, it may then become quite willing to undertake multilateral trade liberalization. Thus, the possibility of preferential trade liberalization amongst others can *induce* a country to participate in multilateral trade liberalization. While we demonstrate the strong building bloc result in the oligopoly model of intraindustry trade (described in greater detail below), it is clear that the mechanism underlying it is not model specific and is likely to arise in most existing models of international trade.

Formally, we analyze the coalition proof Nash equilibria (also called stable equilibria) of two games of trade liberalization between three countries.[2]

[1] See Chang and Winters (2002) for detailed evidence showing that the formation of the Latin American customs union MERCOSUR adversely affected non-member countries by lowering the prices of their exports to MERCOSUR.

[2] Our terminology follows Dutta and Mutuswami (1997).

Under the *FTA game,* countries can pursue either bilateral, multilateral, or no trade liberalization whereas under the *No FTA game,* they have only the latter two options. It is worth noting that under the FTA game each country is free to pursue multiple FTAs – this is important because, on average, each country today belongs to six PTAs (World Bank, 2005). Our underlying framework is one of intraindustry trade under oligopoly where the production cost of the oligopolistic good can differ across countries (see Brander and Krugman, 1983). The FTA game proceeds as follows. In the first stage, each country announces the set of countries with whom it wants to form a bilateral FTA (under which member countries abolish tariffs on each other). An FTA between two countries arises iff they both announce each other's name. Similarly, free trade emerges iff all countries call each other's names. Next, firms compete in the product market in a Cournot fashion. In contrast to the FTA game, in the first stage of the No FTA game, each country can choose between only two alternatives: it can either announce in favor of free trade or no agreement (wherein they retain their existing tariffs on each other). The rest of the No FTA game proceeds just like the FTA game.

Since markets are assumed to be segmented, from each country's perspective, an FTA embodies the following trade-off. On the one hand, forming a bilateral FTA lowers a country's domestic surplus relative to no agreement because of the tariff concession it grants to the other member. On the other hand, the reciprocal tariff reduction granted by the other member increases its export profits. Utilizing this trade-off, we obtain several interesting results. First, we show that under symmetry, free trade is a stable equilibrium of both games. However, while free trade is the unique stable agreement of the No FTA game, a bilateral FTA can *also* be a stable equilibrium of the FTA game (*weak stumbling bloc effect*). Second, there exists patterns of cost asymmetry for which multilateral free trade is an equilibrium *only if* countries have the option to form FTAs (*strong building bloc effect*). Third, despite the absence of any political economy considerations in our model, free trade can fail to be an equilibrium even when FTAs are *not* permissible if one of the countries is sufficiently high cost relative to others. Under such a scenario, FTAs can act as *partial building blocs* by providing welfare improving bilateral trade liberalization. Fourth, a 'hub' and 'spoke' type arrangement – a regime where one country has a bilateral FTA with both its trading partners who in turn do *not* have an FTA with each other – fails to be a stable equilibrium under symmetry. In other words, some degree of cost asymmetry is essential for such a regime to emerge as a stable equilibrium. Finally, it is worth noting that FTAs never act as strong stumbling blocs in our model – i.e. it is never the case that the FTA game yields no agreement whereas the No FTA games yields free trade.

Given the importance of the topic, it is no surprise that there exists a voluminous literature on PTAs. In what follows, we only discuss papers that are closely related to ours and refer the reader to Bhagwati et. al. (1999) for a collection of many of the important contributions in the area. Our paper shares some key elements with Furusawa and Konishi (2007), Goyal and Joshi (2006), Aghion et. al. (2007), and Krishna (1998). Both Furusawa and Konishi (2007) and Goyal and Joshi (2006) apply the network formation game of Jackson and Wolinsky (1996) to examine whether or not a given trade configuration is pairwise stable.[3] Under symmetry, they find that the continued formation of bilateral FTAs leads to global free trade. However, while they examine whether or not bilateralism results in global free trade, they do not consider the consequences of adopting a strictly multilateral approach to global trade negotiations, a comparison that is central to our paper. As a result, they do not address the issue of when and why countries choose to pursue bilateral trade liberalization when multilateral trade liberalization is an option, an issue that lies at the heart of the complex relationship between the two types of liberalization. Using a three-country version of the Brander-Krugman model (also utilized

[3]Relative to our approach, the concept of pairwise stability implies two constraints. First, the deviating coalition can contain at most two countries. Second, a deviation can consist of severing just one existing link or forming one additional link. In order to eliminate these constraints, we follow Dutta and Mutuswami (1997) and use the concept of coalition proof Nash equilibrium to isolate stable equilibria.

by us), Krishna (1998) has shown that the formation of an FTA between two countries reduces their incentives to liberalize trade with respect to the third country. However, in a model with endogenous tariffs, Ornelas (2005a) shows that an FTA induces member countries to lower their tariffs on the non-member country (which in turn reduces its incentive to participate in multilateral trade liberalization). Unlike us, Krishna (1998) and Ornelas (2005a) do not develop an equilibrium theory of FTAs and instead consider the effects of an exogenously given FTA.[4]

Our conceptual approach is related to that of Aghion et. al. (2007) who examine a leading country's choice between sequential and multilateral bargaining of free trade agreements. Like us, Aghion et. al. (2007) also identify building and stumbling bloc effects of FTAs. However, there are important differences between their approach and ours. First, in our model, *all* countries are free to negotiate FTAs and not just a single leading country.[5] Also, countries are free to form a pair of bilateral FTAs in our model and are not required to choose between joining a single grand coalition or staying out. Second, our analysis complements theirs in two important respects (*i*) we assume that governments care aggregate social welfare whereas their examples illustrating the effects of FTAs assume governments care only about producer surplus and (*ii*) unlike them but like Grossman and Helpman (1995), we do not allow transfers between different coalitions.[6] Point (*ii*) is important because when transfers are possible and

[4] Ornelas (2005b) provides an analysis of political economy considerations that arise in the context of FTAs.

[5] Aghion et. al. (2007) do consider extensions where the leadership role is assigned to other countries if the first leader's offer is not accepted by the followers but they focus on deriving necessary conditions for a free trade equilibrium.

[6] Grossman and Helpman (1995) point out that transfers are rarely used in trade agreements and when used they are limited in scope. However, Aghion et. al. (2007) note that the exchange of concessions on non-trade-related issues among FTA members can be viewed as transfers. Both arguments capture certain aspects of reality. Since one of our goals is to provide a non-cooperative theory of FTA formation in a game with non-transferable utility, we proceed with the assumption of no transfers. In this context, it is noteworthy that Raimondos-Møller and Woodland (2006) have shown that if non-discriminatory tariff reforms by a subset of countries are accompanied by appropriate income transfers between them, reforming members can make themselves strictly better

there is grand coalition superadditivity, in the absence of externalities free trade necessarily emerges in equilibrium regardless of whether the leading country chooses a sequential or multilateral approach.[7] In our model, even when free trade is Pareto optimal (as it is under symmetry), a bilateral FTA can emerge in equilibrium.

Our model is also related to that of Riezman (1999) who asks whether the option to pursue FTAs facilitates or hinders the achievement of free trade. However, while we analytically derive the coalition proof Nash equilibria of two non-cooperative games, 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 between countries in a way that cannot be done in the inter-industry trade framework utilized by Riezman (1999). As has already been noted, cost asymmetry between countries plays a crucial role in determining conditions under FTAs act as partial and strong building blocs. In this context, it is worth noting that both Krugman (1991) and Grossman and Helpman (1995) pointed out that asymmetries across countries could play an important role in determining the relationship between bilateral and multilateral trade liberalization.

The effects of free trade agreements have also been analyzed in models of repeated interaction between countries that require multilateral cooperation to be self-enforcing – see Bagwell and Staiger (1997), Bond and Syropoulos (1996), Conconi and Perroni (2003), Freund (2000), and Saggi (2006). We add value to this literature by treating both bilateral and multilateral liberalization as endogenous. Maggi and Rodriguez-Clare (1998) provide a completely different perspective on international trade agreements by showing that, in the presence of domestic protectionist pressures, such agreements

off without having an adverse effect on non-members.

[7] Grand coalition superadditivity holds if the *joint* payoff of the three countries is larger under free trade than under no FTAs whatsoever or a bilateral FTA between any two countries. When this condition fails, Aghion et. al. (2007) show that the nature of externalities created by FTAs assumes a crucial role: when such externalities are negative, FTAs necessarily facilitate the achievement of global free trade whereas when they are positive, they hamper it. In our model, a bilateral FTA necessarily generates a negative externality for the non-member.

can improve the domestic allocation of resources by helping a government credibly commit to free trade.[8]

2 Model

There are three countries (a, b, c) and two goods: x and y. Preferences over the two goods are quasilinear: $U(x, y) = u(x) + y$. Good x is produced by a single profit-maximizing firm in each country at a constant marginal cost in terms of the numeraire good y.[9] Firms compete in quantities and make independent decisions regarding how much to sell in each market (i.e. markets are segmented as in Brander and Krugman, 1983).

2.1 Production and trade

Due to market segmentation, it is sufficient to focus on only one country's market. Let t denote a country's tariff on a trading partner with whom it does not have an FTA. In other words, under no agreement (i.e. the status quo) each country imposes the tariff t on both its trading partners whereas if it has an FTA with both of them it practises free trade. As Grossman and Helpman (1995) note, GATT Article XXIV forbids FTA members from raising their tariffs on non-members. Accordingly, like Duttagupta and Panagariya (2006) we assume that FTA members retain their status quo tariff t on the non-member.

We now describe production and trade under no agreement. Firm j's effective marginal cost of exporting equals $\zeta_j + t$ where $\zeta_j \geq 0$ equals its marginal cost of production for good x. By assumption, countries impose no taxes on local firms and the numeraire good that may be traded internationally in order to balance trade.

[8] In a recent paper, Maggi and Rodriguez-Clare (2007) build a model of trade agreements that integrates the terms of trade motive for FTAs emphasized by Bagwell and Staiger (1997) with the commitment argument analyzed in Maggi and Rodriguez-Clare (1998).

[9] The monopoly assumption is not necessary. We just need firms to have market power in order to justify the existence of tariffs in a welfare-maximizing framework.

Let x_{ji} denote country j's exports to country i; x_{ii} the sales of firm i in country i; and $x_i = x_{ii} + \sum_j x_{ji}$ denote total sales of good x in country i. Country j's profit function for exports to country i, denoted by Π_{ji}, can be written as:

$$\Pi_{ji} = [p_i(x_i) - \zeta_j - t]x_{ji} \tag{1}$$

First order conditions (FOCs) for profit maximization for exporters are

$$p_i + p_i' x_{ji} = \zeta_j + t \tag{2}$$

The above FOCs together with an analogous condition for the local firm (i.e. $p_i + p_i' x_{ii} = \zeta_i$) determine the equilibrium output levels of all firms. Summing the FOCs for all firms in country i gives

$$3p_i + p_i' x_i = 2t + \sum_z \zeta_z \text{ where } z = a, b, c. \tag{3}$$

Following Bergstrom and Varian (1985), if the left hand side is decreasing in x_i (i.e. $4p_i' + p_i'' x_i < 0$) then total industry output sold in country i depends only on the sum of the (tariff included) marginal costs of production of all firms. Assume this property holds so that we have:

$$\frac{dx_i}{dt} < 0 \tag{4}$$

In addition, the following comparative statics are also assumed to hold:[10]

$$\frac{dx_{ji}}{dt} < 0 < \frac{dx_{ii}}{dt} \tag{5}$$

In other words, an increase in the tariff rate (t) lowers country j's exports to county i (x_{ji}) while it increases the sales of its local firm (x_{ii}).

Welfare of country i is defined as the sum of its domestic surplus and total export profits:

$$W_i \equiv S_i + \sum_{j \neq i} \Pi_{ij} \tag{6}$$

[10] As is well known, if second order conditions for profit maximization hold these comparative statics obtain when the Cournot Nash equilibrium is stable and output levels of firms are strategic substitutes. Since these results are well known, it is convenient to directly assume that the comparative statics in (5) hold.

where domestic surplus S_i is defined as

$$S_i \equiv u(x_i) - p_i x_i + \Pi_{ii} + t \sum_j x_{ji} \tag{7}$$

where $u(x_i) - p_i x_i$ is consumer surplus in country i; $\Pi_{ii} = (p_i - \zeta_i)x_{ii}$ equals firm i's profits in its own market; $t \sum_j x_{ji}$ equals country i's tariff revenue; and $\Pi_{ij} = (p_i - \zeta_i - t)x_{ij}$ its profits in foreign market j, where $j \neq i$. World welfare is defined the sum of the welfare of individual countries:

$$WW = W_i + W_j + W_k \tag{8}$$

Since the formation of FTAs results in the elimination of some of the tariffs, in what follows in functions $S(.)$ and $W(.)$ we list the tariffs faced by foreign countries in ascending alphabetical order. Also, in the export profit function $\Pi_{ij}(.)$, the first argument is the tariff faced by country i while the second argument is the tariff faced by its rival exporter (i.e. country k).

As is well known from the work of Brander and Spencer (1984), under fairly general conditions each country has a unilateral incentive to impose rent extracting tariffs on its trading partners (unless it commits not to do so via an FTA). Accordingly, we assume that the following holds:

Assumption 1:

$$S_i(t,t) > \max\{S_i(0,t),\ S_i(t,0)\} > S_i(0,0)$$

i.e. domestic surplus of each country is highest under no agreement and lowest when it practises free trade. Since countries are asymmetric, a comparison of $S_i(0,t)$ and $S_i(t,0)$ is ambiguous in general.

Let $\mathcal{T}$ be the set of all feasible trade policy regimes:

$$\mathcal{T} = \{\langle\{\Phi\}\rangle, \langle\{ab\}\rangle, \langle\{ac\}\rangle, \langle\{bc\}\rangle, \langle\{ab, ac\}\rangle, \langle\{ab, bc\}\rangle, \langle\{ac, bc\}\rangle, \text{ or } \langle\{F\}\rangle\}$$

where $\langle\{\Phi\}\rangle$ denotes the status quo and and $\langle\{F\}\rangle$ denotes global free trade. Let r and v be any two elements of $\mathcal{T}$. Further, let $w_i(r)$ denote country i's welfare and $ww(r)$ denote world welfare under regime r. Note that small letters denote functions of trade regimes whereas capital letters denote functions of tariffs that prevail during those regimes. For example, country i's welfare under the FTA $\langle\{ij\}\rangle$ can be written either as $w_i(ij)$ or as $W_i(0,t)$. Similar notation applies to domestic surplus and export profit functions.

Define $\Delta w_i(r-v)$ as the difference between country i's welfare under regimes r and v:

$$\Delta w_i(r-v) \equiv w_i(r) - w_i(v) \tag{9}$$

As is well known, under Cournot competition, the higher a country's cost of producing good x, the smaller its volume of exports and the larger its volume of imports (of this good). Given this and Grossman and Helpman's (1995) argument that an FTA is more likely to obtain when trade between potential partners is relatively balanced, we make the following assumption:

Assumption 2:

$$\frac{\partial \Delta w_i(r-v)}{\partial \zeta_i} \leq 0 \leq \frac{\partial \Delta w_i(r-v)}{\partial \zeta_m}$$

where m is an FTA partner of country i under regime r (but not regime v).

To get further insight behind assumption 2, consider regimes $\langle\{\Phi\}\rangle$ and $\langle\{ij\}\rangle$ from country i's perspective. The intuition underlying $\frac{\partial \Delta w_i(ij-\Phi)}{\partial \zeta_i} \leq 0$ is as follows. Because of their larger volume of imports, higher cost countries have relatively *more to gain* from using tariffs. Similarly, due to the smaller volume of their exports, higher cost countries have *less to lose* from other countries' tariffs. As a result, a country's willingness to enter into a bilateral FTA with another depends negatively on its own cost.

A similar intuition underlies $\frac{\partial \Delta w_i(ij-\Phi)}{\partial \zeta_j} \geq 0$. The higher the production cost of its trading partner, the larger the increase in export profits enjoyed by a country due to the trade liberalization undertaken by its partner and the smaller the loss in local profits suffered by the domestic firm due to its own trade liberalization.

3 Endogenous Free Trade Agreements

Consider the following two stage game of bilateral trade liberalization (called the FTA game). In the first stage, each country 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). Next, given tariffs, firms compete in product markets. As is clear from the structure of the FTA game, an FTA member can sign an independent FTA with the non-member *without* needing consent of the other member.

A country's strategy set consists of four possible announcements. Country i's strategy set Ω_i^F is:

$$\Omega_i^F = \{\{\phi, \phi\}, \{j, \phi\}, \{\phi, k\}, \{j, k\}\} \tag{10}$$

where $\{\phi, \phi\}$ is an announcement in favor of no agreement with either of its trade partners. In order to conserve notation, each trade policy regime is denoted as follows: (i) No agreement $\langle\{\Phi\}\rangle$ is maintained 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 they both announce each other's name $j\epsilon\alpha_i$ and $i\epsilon\alpha_j$; (iii) two independent bilateral FTAs $\langle\{ij, ik\}\rangle$ in which i is the common member 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 $\langle\{F\}\rangle$ obtains iff all countries announce each others' names: i.e. $\alpha_a = \{b, c\}$, $\alpha_b = \{a, c\}$, and $\alpha_c = \{a, b\}$.

It is worth noting here that the regime under which there exist two independent bilateral FTAs (i.e. $\langle\{ij, ik\}\rangle$) can be viewed as a 'hub and spoke' trading arrangement where the common member (i.e. country i) is the hub while each of the other two countries (i.e. countries j and k) is a spoke. Also note that two different strategy vectors may yield the same agreement(s) when countries' announcements do not match. For example, consider the following announcements:

$$\alpha_a = \{b, \phi\},\ \alpha_b = \{a, c\},\ \alpha_c = \{\phi, b\} \tag{11}$$

The above strategy vector gives rise to two independent FTAs $\langle\{ab, bc\}\rangle$ of which country b is the common member. But the same outcome obtains when the strategy vector is given by:

$$\alpha_a = \{b, c\},\ \alpha_b = \{a, c\},\ \alpha_c = \{\phi, b\} \tag{12}$$

Here, even though country a announces country c, country c wants to form an FTA only with country b.[11] In order to eliminate redundant announcements, assume that each FTA announcement costs ε (where $\varepsilon > 0$ is arbitrarily small).

Our method of analysis is to compare the FTA game with the following game of multilateral trade liberalization (called the No FTA game). In the first stage of the No FTA game, each country announces either in favor of or against free trade. If all countries announce in favor, free trade emerges. If not, the status quo prevails. Next, firms compete in the product market. Clearly, the No FTA game restricts the strategy set of country i to $\Omega_i = \{\{\phi, \phi\}, \{j, k\}\}$, $j \neq k \neq i$. We compare the equilibria of these two games of trade liberalization to determine how the freedom to pursue bilateral FTAs affects the likelihood of obtaining global free trade.

As might be expected, both games admit multiple Nash equilibria. To deal with this multiplicity and to capture the process of FTA formation in a more realistic fashion, we focus attention on Nash equilibria that are immune to *self-enforcing* coalitional deviations (i.e. are coalition proof or *stable*).[12] It is useful to note that in the No FTA game, a unilateral deviation from free trade by any country reverts everyone back to no agreement whereas in the FTA game the same deviation results in the deviating country becoming either (*i*) a non-member country under an FTA between the other two countries or (*ii*) a spoke under a pair of bilateral FTAs. Since the welfare of a country under these trade policy regimes can be lower/higher

[11] Note that under our approach, the formation of a bilateral FTA requires consent from both sides. This is in contrast to the open membership rule analyzed by Yi (1996) where existing members cannot prevent others from joining.

[12] See Bernheim et. al. (1987) for the formal definition of a coalition proof Nash equilibrium.

than its welfare under no agreement, it is *not* immediately obvious under which game the unilateral incentive to deviate from free trade is stronger. Thus, even though the set of possible deviations from free trade under the No FTA game is a strict subset of those under the FTA game, it does *not* follow that free trade is more likely to be a stable equilibrium of the No FTA game.

The analysis proceeds as follows. First, we show that trade liberalization of any kind is desirable under symmetry so that free trade yields higher world welfare than any other policy regime (Lemma 1). We then derive Nash equilibria of the two games (Proposition 1 and 2). Next, we show that when countries are symmetric, free trade is the unique stable equilibrium of the No FTA game (Proposition 3) whereas the FTA game also admits a bilateral FTA as a stable equilibrium (Proposition 4). Next, we show that free trade fails to be a stable equilibrium of the FTA game if one of the countries is sufficiently high cost relative to the other two (Proposition 5). Continuing with the same underlying asymmetry, we then provide conditions under which free trade is a stable equilibrium of the FTA game (Proposition 7) but *not* that of the No FTA game (Corollaries 1 and 2). Propositions 8 and 9 describe conditions under which different patterns of FTAs emerge as stable equilibria. Finally, for the case of linear demand, we graphically illustrate stable equilibria of the two games; isolate the building and stumbling bloc effects of FTAs; and examine how the FTA option affects the welfare of individual countries and the world as a whole (Proposition 10).

4 Equilibrium FTAs under symmetry

Through-out this section, we assume that the cost of producing good x is equal across countries: $\zeta_i = \zeta$ for all i. It proves convenient to begin with the effects of FTAs on global welfare.

4.1 Trade liberalization and welfare

We show in the appendix that a reduction in any country's tariff(s) (whether on a preferential or a non-discriminatory basis) increases aggregate world welfare. This implies the following:

Lemma 1: *Under symmetry, free trade yields higher world welfare than any other trade policy regime:* $ww(F) > ww(ij, ik) > ww(ij) > ww(\Phi)$.

The intuition behind this result is simple: when all countries have the same cost of production, the allocation of output across countries is immaterial and any trade restrictions (whether preferential or multilateral) simply lower aggregate world output and therefore welfare. Since any tariff creates a deadweight loss, rent extraction by an importing country is more than offset by the loss in profits of exporters.

4.2 Nash equilibria

Before deriving (subgame perfect) Nash equilibria of the two games, 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 $\langle\{\Phi\}\rangle$.

It is clear that no agreement $\langle\{\Phi\}\rangle$ is a Nash equilibrium of the No FTA game since no country has a unilateral incentive to announce another country's name if the latter does not announce its name in return. Furthermore, symmetry implies that under no agreement $\langle\{\Phi\}\rangle$, welfare of all countries is equal. Clearly, the same is true under free trade $\langle\{F\}\rangle$. Since world welfare is higher under $\langle\{F\}\rangle$ than under $\langle\{\Phi\}\rangle$ (Lemma 1), it follows that each country is better off under $\langle\{F\}\rangle$ than under $\langle\{\Phi\}\rangle$. As a result, under symmetry no country has a unilateral incentive to deviate from free trade since *any* other announcement on its part leads to no agreement where it (and everyone else) is worse off:

$$w_i(\Phi) < w_i(F) \text{ for all } i \tag{13}$$

Proposition 1: *Under symmetry, no agreement* $\langle\{\Phi\}\rangle$ *and free trade* $\langle\{F\}\rangle$ *are both Nash equilibria of the No FTA game.*

The following two conditions prove useful in describing the Nash equilibria of the FTA game:

Condition 1:

$$w_k(ij, ik) < w_k(ij) \tag{14}$$

i.e. if condition 1 holds, each country prefers to be a non-member under a bilateral FTA to being a spoke under a pair of bilateral FTAs.

Condition 1 can be understood as follows. The tariff reduction that country k receives from country i under $\langle\{ij, ik\}\rangle$ removes the disadvantage it faces relative to country j while exporting to country i's market under $\langle\{ij\}\rangle$. However, to achieve equal footing with country j in country i's market, country k has to grant preferential access to country i in its hitherto fully protected market. Thus, country k's preference among regimes $\langle\{ij\}\rangle$ and $\langle\{ij, ik\}\rangle$ is ambiguous in general.

Condition 2:

$$w_i(F) < w_i(ij) \tag{15}$$

i.e. if condition 2 holds, each country is better off as a member of a bilateral FTA relative to free trade – i.e. two countries find it beneficial to exclude the third.

It is clear that no agreement $\langle\{\Phi\}\rangle$ is always a Nash equilibrium of the FTA game. Is a bilateral FTA $\langle\{ij\}\rangle$ a Nash equilibrium too? Comparing country k's welfare under $\langle\{ij\}\rangle$ and $\langle\{\Phi\}\rangle$, we note that export profits of country k are higher under $\langle\{\Phi\}\rangle$ relative to $\langle\{ij\}\rangle$ (i.e. $\Pi_{ki}(t,t) > \Pi_{ki}(t,0)$ and $\Pi_{kj}(t,t) > \Pi_{kj}(t,0)$) whereas its domestic surplus under the two regimes is the same (i.e. $s_k(ij) = s_k(\Phi) = S_k(t,t)$). Therefore, we have the following result:

Lemma 2: *A bilateral FTA between countries i and j makes country k worse off relative to no agreement:*

$$w_k(ij) < w_k(\Phi) \tag{16}$$

Since world welfare is higher under $\langle\{ij\}\rangle$ relative to $\langle\{\Phi\}\rangle$, the above inequality implies that the sum of countries i and j's welfare must surely be higher under $\langle\{ij\}\rangle$:

$$w_i(ij) + w_j(ij) > w_i(\Phi) + w_j(\Phi) \tag{17}$$

Since countries i and j are symmetric, we must have

$$w_i(ij) = w_j(ij) > w_i(\Phi) = w_j(\Phi) \tag{18}$$

Hence, a member country of a bilateral FTA has no incentive to deviate to no agreement and a bilateral FTA $\langle\{ij\}\rangle$ is indeed a Nash equilibrium of the FTA game.

Is a hub and spoke arrangement such as $\langle\{ij, ik\}\rangle$ a Nash equilibrium? Before addressing this question we note that

$$w_i(ij, ik) > w_i(F) \tag{19}$$

The logic behind this inequality is as follows. Starting at free trade, if country k revokes its FTA with country j, export profits of country i increase in both markets because its rival exporters face tariffs whereas it itself does not: $\Pi_{ij}(0,t) > \Pi_{ij}(0,0)$ and $\Pi_{ik}(0,t) > \Pi_{ik}(0,0)$. Furthermore, the domestic surplus of country i does not change relative to free trade since its own tariff equals zero under both regimes: $s_i(ij, ik) = s_i(F) = S_i(0,0)$. As a result, country i's welfare under $\langle\{ij, ik\}\rangle$ is *higher* than that under $\langle\{F\}\rangle$.

Furthermore, we assert that $w_j(ij, ik) < w_j(ij)$: the move from $\langle\{ij\}\rangle$ to $\langle\{ij, ik\}\rangle$ makes country j worse off since it loses its preferential status in country i's market. Next, note from Lemma 1 that $ww(ij, ik) > ww(ij)$. As a result, either country i or country k or both of them are better off under $\langle\{ij, ik\}\rangle$ relative to $\langle\{ij\}\rangle$. Given the fact that the hub country's welfare under $\langle\{ij, ik\}\rangle$ exceeds even that under free trade (see 19), we make the following assumption:

Assumption 3: *The hub country of a pair of independent FTAs enjoys higher welfare than a member country of a single FTA:*

$$w_i(ij, ik) > w_i(ij) = w_i(ik) \tag{20}$$

We are now ready to investigate whether $\langle\{ij, ik\}\rangle$ is a Nash equilibrium. Three possible unilateral deviations from $\langle\{ij, ik\}\rangle$ need to be considered:

- UP1: Country i's deviation from $\langle\{ij, ik\}\rangle$ to $\langle\{\Phi\}\rangle$.
- UP2: Country i's deviation from $\langle\{ij, ik\}\rangle$ to $\langle\{ij\}\rangle$ (or $\langle\{ik\}\rangle$).
- UP3: Country k's deviation from $\langle\{ij, ik\}\rangle$ to $\langle\{ij\}\rangle$.

Deviation UP1 can be ruled out because $w_i(ij, ik) > w_i(\Phi)$ whereas deviation UP2 can be ruled out on the basis of Assumption 3. Finally, note that deviation UP3 cannot occur if condition 1 fails: $w_k(ij, ik) > w_k(ij)$. As a result, a pair of bilateral FTAs $\langle\{ij, ik\}\rangle$ is a Nash equilibrium if condition 1 fails.

Is free trade also a Nash equilibrium of the FTA game? To be able to answer this question in the affirmative, we need to rule out the following two deviations:

- UF1: Country k's deviation from $\langle\{F\}\rangle$ to $\langle\{ij\}\rangle$.
- UF2: Country k's deviation from $\langle\{F\}\rangle$ to $\langle\{ij, ik\}\rangle$ (or $\langle\{ij, jk\}\rangle$).

Inequalities (13) and (16) imply that deviation UF1 cannot occur:

$$w_k(F) > w_k(\Phi) > w_k(ij) \tag{21}$$

Furthermore, deviation UF2 can be ruled out due to the following result:

Lemma 3: *Under the pair of independent bilateral FTAs $\langle\{ij, ik\}\rangle$, each spoke country (i.e. j and k) is worse off relative to free trade.*

The logic behind this result is as follows. Since world welfare is lower under $\langle\{ij, ik\}\rangle$ relative to free trade $\langle\{F\}\rangle$ whereas the welfare of the hub country is higher (see inequality 19), the sum of the welfare of the two spoke countries (i.e. j and k) must be *lower* than that under free trade. Since both are symmetric, it follows that both must be worse off under $\langle\{ij, ik\}\rangle$ relative to free trade $\langle\{F\}\rangle$:

$$w_j(ij, ik) = w_k(ij, ik) < w_j(F) = w_k(F) \tag{22}$$

The following proposition summarizes the Nash equilibria of the FTA game:

Proposition 2: *No agreement $\langle\{\Phi\}\rangle$, a bilateral FTA $\langle\{ij\}\rangle$, and free trade $\langle\{F\}\rangle$ are all Nash equilibria of the FTA game. In addition, if condition 1 fails then a pair of bilateral FTAs $\langle\{ij, ik\}\rangle$ is also a Nash equilibrium.*

Propositions 1 and 2 show that both games admit multiple Nash equilibria. To resolve this multiplicity (as well as to capture the process of FTA formation in a richer fashion), we now focus attention on Nash equilibria that are immune to self-enforcing coalitional deviations (i.e. are coalition proof or *stable*).

4.3 Stable Nash equilibria

We begin with the No FTA game. Recall from Proposition 1 that the No FTA game admits two Nash equilibria: $\langle\{\Phi\}\rangle$ and $\langle\{F\}\rangle$. Which, if any, of these is stable? It is easy to see that all three countries have a joint incentive to deviate from $\langle\{\Phi\}\rangle$ to $\langle\{F\}\rangle$ – each is better off under free trade than under no agreement. Based on the definition of a stable Nash equilibrium, $\langle\{\Phi\}\rangle$ fails to be stable if the initial deviation from $\langle\{\Phi\}\rangle$ to $\langle\{F\}\rangle$ is self-enforcing – i.e. no country or a pair of countries has an incentive to *further deviate* from $\langle\{F\}\rangle$. This indeed is the case because any such deviation (unilateral or coalitional) reverts the world back to status quo under which everyone is worse off. As a result, we have the following:

Proposition 3: *Free trade is the unique stable equilibrium of the No FTA game.*

This result implies that when countries are symmetric, FTAs cannot *possibly* act as building blocs. We now turn to the FTA game and examine which Nash equilibria are stable.

We begin with no agreement $\langle\{\Phi\}\rangle$. Consider a member country's welfare under the FTA $\langle\{ij\}\rangle$ relative to no agreement $\langle\{\Phi\}\rangle$. It is immediate from (18) that countries i and j always have an incentive to jointly deviate from $\langle\{\Phi\}\rangle$ to a bilateral FTA $\langle\{ij\}\rangle$. Furthermore, since $\langle\{ij\}\rangle$ is a Nash equilibrium of the FTA game (Proposition 2), there can be no unilateral deviations from it. As a result, the initial joint deviation of countries i and j from $\langle\{\Phi\}\rangle$ to a bilateral FTA $\langle\{ij\}\rangle$ is self-enforcing and no agreement $\langle\{\Phi\}\rangle$ is *not* stable.

Now consider a hub and spoke arrangement such as $\langle\{ij, ik\}\rangle$ as a candidate for a stable Nash equilibrium. We know from inequality (22) that countries j and k always benefit from a joint deviation from $\langle\{ij, ik\}\rangle$ to global free trade $\langle\{F\}\rangle$. Once again, this joint deviation is self-enforcing since $\langle\{F\}\rangle$ is a Nash equilibrium. As a result, a pair of bilateral FTAs $\langle\{ij, ik\}\rangle$ is also *not* stable.

Two candidates for stable equilibria remain: free trade $\langle\{F\}\rangle$ and a bilateral FTA $\langle\{ij\}\rangle$. Consider free trade first. For free trade to be stable, we need to rule out three types of coalitional deviations by a pair of countries:

- 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 $\langle\{ij\}\rangle$.

Since $w_i(F) > w_i(\Phi)$(inequality 13) the joint deviation JF1 cannot occur. Similarly, since $w_j(ij, ik) < w_j(F)$ (inequality 22), we can rule out deviation JF2. Can deviation JF3 be ruled out? Clearly, if country i is worse off under a bilateral FTA relative to free trade (i.e. condition 2 fails) then deviation JF3 cannot occur. Under such a scenario, free trade $\langle\{F\}\rangle$ is immune to all coalitional (as well as unilateral) deviations and is in fact the unique stable equilibrium of the FTA game.

But what if countries i and j have an incentive to jointly deviate from free trade $\langle\{F\}\rangle$ to a bilateral FTA $\langle\{ij\}\rangle$ (i.e. condition 2 holds)? To determine whether this deviation is self-enforcing or not, we need to consider two further deviations from $\langle\{ij\}\rangle$:

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

It is immediate from Lemma 1 and inequality (18) (i.e. $w_i(ij) > w_i(\Phi)$) that deviation FD1 cannot occur. What about deviation FD2? From assumption 3 (see inequality 20) we know that, taking country k's announcement as given, country i has an incentive to further deviate from $\langle\{ij\}\rangle$ to the pair of bilateral FTAs $\langle\{ij, ik\}\rangle$ – i.e. deviation FD2 *will indeed occur*. As a result, the initial joint deviation of countries i and j from free trade to the bilateral FTA $\langle\{ij\}\rangle$ (i.e. deviation JF3) is *not* self-enforcing. Thus, free trade $\langle\{F\}\rangle$ is stable even when condition 2 holds.

Can $\langle\{ij\}\rangle$ also be stable when condition 2 holds? It turns out that the answer to this question depends on whether condition 1 holds or not. First consider the case where condition 1 fails – i.e. countries i and k have a joint incentive to deviate from $\langle\{ij\}\rangle$ to $\langle\{ij, ik\}\rangle$. We know from Proposition 2 that $\langle\{ij, ik\}\rangle$ is a Nash equilibrium when condition 1 fails. Therefore, when condition 1 fails the initial deviation of countries i and k from $\langle\{ij\}\rangle$ to $\langle\{ij, ik\}\rangle$ is self-enforcing. As a result, if condition 2 holds and condition 1 fails, bilateral FTA $\langle\{ij\}\rangle$ is *not* stable.

Now consider the case where condition 1 holds. First, since condition 2 holds, there exists no incentive for all countries to multilaterally deviate from $\langle\{ij\}\rangle$ to free trade $\langle\{F\}\rangle$. Second, since condition 1 holds, countries i and k have no joint incentive to deviate from $\langle\{ij\}\rangle$ to $\langle\{ij, ik\}\rangle$. Finally, since condition 2 holds, inequality (22) implies that country i has no incentive to jointly deviate with countries j and k from $\langle\{ij\}\rangle$ to $\langle\{ij, jk\}\rangle$ (or $\langle\{ik, jk\}\rangle$). As a result, a bilateral FTA $\langle\{ij\}\rangle$ is stable if both condition 1 and condition 2 hold. We summarize as follows:

Proposition 4: *Free Trade* $\langle\{F\}\rangle$ *is a stable Nash equilibrium of the FTA game and it is uniquely stable if either condition 1 or condition 2 fail. A bilateral FTA* $\langle\{ij\}\rangle$ *is a stable Nash if both condition 1 and condition 2 hold. Finally, no agreement* $\langle\{\Phi\}\rangle$ *and a pair of bilateral FTAs* $\langle\{ij, ik\}\rangle$ *are not stable Nash equilibria.*

The above proposition clarifies that under symmetry FTAs end up acting as *weak stumbling blocs*: while free trade is uniquely stable under the No FTA game, there exist conditions under which a bilateral FTA is also stable under the FTA game. Yet, this is not a particularly strong indictment of FTAs – while free trade loses its uniqueness, it still continues to be a stable equilibrium.

Given these results, it is natural to ask whether a different role for FTAs might emerge when countries are not necessarily symmetric (say with respect to their costs of production of good x). For example, is it possible that in the absence of symmetry, two countries are willing to enter into a bilateral FTA but unwilling to engage in multilateral free trade? Even more interestingly, can the option to form bilateral FTAs facilitate the obtainment of multilateral free trade? We now turn to these questions.

5 FTAs among asymmetric countries

From hereon, we drop the assumption that the production cost of good x is equal across countries. It proves instructive to focus on the case where two countries have symmetric and low costs relative to the third. Accordingly, throughout the analysis under asymmetry, let $\zeta_a = \zeta_b = 0$ and $\zeta_c = \zeta > 0$. Note also that Lemma 1 requires a slight modification: trade liberalization necessarily improves world welfare under asymmetry as long as it is not biased against low cost producers. In other words we know the following:[13]

$$ww(F) > ww(ab, ac) > ww(ab) > \max\{ww(\Phi), ww(ac)\} \tag{23}$$

[13] Note that $ww(ab, ac) = ww(ab, bc)$ and $ww(ac) = ww(bc)$.

Intuitively, a bilateral FTA between a low cost and a high cost country has two conflicting effects on world welfare. On the one hand, it increases world welfare by raising the aggregate output of good x. On the other hand, it diverts production away from a low cost source to a high cost one and this adverse allocation effect harms world welfare. The proof of Lemma 1 can be modified to show that as long as the asymmetry between countries is not too high, an FTA between a high and low cost country also increases world welfare (since the allocation distortion is mild in such circumstances).

We begin with the No FTA game and first show that under cost asymmetry global free trade may fail to obtain even when countries are not free to form bilateral FTAs.

5.1 Feasibility of free trade

Consider the perspective of the two low cost countries (denoted by $i = a$, b). From proposition 1 we know that under symmetry, the welfare of country a low cost under free trade $\langle\{F\}\rangle$ is higher than that under no agreement $\langle\{\Phi\}\rangle$:

$$\lim_{\zeta \longrightarrow 0} \Delta w_i(F - \Phi) > 0 \text{ for } i = a, b \tag{24}$$

Assumption 2 implies:

$$\frac{\partial \Delta w_i(F - \Phi)}{\partial \zeta} > 0 \text{ for } i = a, b \tag{25}$$

i.e. the higher the cost of country k, the larger the gains of multilateral trade liberalization for the low cost countries.

Inequalities (24) and (25) together imply that the two low cost countries have no incentive (joint or unilateral) to deviate from free trade to no agreement:

$$w_i(F) > w_i(\Phi) \text{ for } i = a, b \tag{26}$$

Next, consider the high cost country's (i.e. c's) perspective. Define ζ^P to be the prohibitive cost level at which the export profits of country c under free trade equal zero in each foreign market:

$$\lim_{\zeta \longrightarrow \zeta^P} \pi_{ca}(F) = \lim_{\zeta \longrightarrow \zeta^P} \pi_{cb}(F) = 0 \tag{27}$$

Since domestic surplus of each country is higher under no agreement $\langle\{\Phi\}\rangle$ than under free trade $\langle\{F\}\rangle$ the following is immediate:

$$\lim_{\zeta \longrightarrow \zeta^P} \Delta w_c(\Phi - F) > 0 \text{ since } s_c(\Phi) > s_c(F) \tag{28}$$

Inequalities (24), (28), and assumption 2 imply that there exists a critical threshold cost level (ζ^{ϕ}) such that:

$$w_c(\Phi) \geq w_c(F) \text{ iff } \zeta \geq \zeta^{\phi} \tag{29}$$

Intuitively, opening up its market is unattractive to the high cost country when it stands to gain very little (or nothing) from foreign trade liberalization. Thus, the following obtains under asymmetry:

Proposition 5: *Free trade is the unique stable equilibrium of the No FTA game iff $\zeta < \zeta^{\phi}$.*[14] *Or else, no agreement obtains.*

Thus, global free trade fails to be a stable equilibrium even when countries lack the option to form bilateral FTAs as long as one of them is sufficiently high cost relative to the others (i.e. $\zeta > \zeta^{\phi}$). Recall from Proposition 3 that under symmetry free trade is uniquely stable under the No FTA game and when such is the case, FTAs can only act as stumbling blocs. Is this also true when countries are asymmetric? We now show that this is not so.

5.2 FTAs as strong building blocs

In this sub-section we show that the option to form bilateral FTAs can serve as *strong building blocs* – there exists circumstances where free trade is stable under the FTA game whereas it is not so under the No FTA game. Intuitively, this result obtains since the high cost country prefers no agreement to free trade which in turn it prefers to a bilateral FTA between the other two countries: $w_c(\Phi) > w_c(F) > w_c(ab)$. The key point is that under a purely multilateral approach, the high cost country can ensure the preservation of the status quo by voting against free trade whereas it cannot do

[14] When $\zeta = \zeta^{\phi}$, both no agreement and free trade are stable. An analogous statement applies to the remainder of the propositions under asymmetry.

so when bilateral FTAs are permitted. As a result, in the FTA game it can end up calling in favor of free trade to avoid ending up as a non-member with the other two countries in a FTA. Before showing this, we state the following result (proved in the appendix):

Proposition 6: *There exist no unilateral or self-enforcing joint deviations involving the two low cost countries from free trade.*

Given this result, it is clear that the viability of free trade depends critically upon the preferences of the high cost country. Following the definition of ζ^{ϕ} (in 29) let ζ^{r} define the critical threshold cost level below which the high cost country prefers free trade to regime r:

$$w_c(F) \geq w_c(r) \text{ iff } \zeta \leq \zeta^{r} \tag{30}$$

where $r = \{\langle\{\Phi\}\rangle, \langle\{ab\}\rangle, \langle\{ac\}\rangle, \langle\{bc\}\rangle, \langle\{ab, ac\}\rangle, \text{or } \langle\{ab, bc\}\rangle\}$. Arguments analogous to those that underlie the existence of ζ^{ϕ} ensure that these critical cost thresholds also exist for the other trade policy regimes.[15]

We know from (16) that the high cost country always prefers no agreement to being a non-member: $w_c(\Phi) > w_c(ab)$. This is because its export profits are always higher under no agreement $\langle\{\Phi\}\rangle$ relative to the bilateral FTA $\langle\{ab\}\rangle$ while its domestic surplus under the two regimes is the same. As a result, we must have $\Delta w_c(\Phi - F) > \Delta w_c(ab - F)$ which implies that the critical cost threshold (ζ^{ϕ}) at which $\Delta w_c(\Phi - F) = 0$ is smaller than the one (ζ^{ab}) where $\Delta w_c(ab - F) = 0$:

Lemma 4: $\zeta^{\phi} < \zeta^{ab}$.

Similarly, by definition, the high cost country has an incentive to unilaterally deviate from free trade $\langle\{F\}\rangle$ to $\langle\{ab, ac\}\rangle$ iff $\zeta > \zeta^{ab,ac}$.[16] Before proceeding further, it is convenient to write down a condition analogous to condition 1:

[15] Note that the threshold ζ^{ac} (similarly ζ^{bc}) exists only when $w_c(ac) < w_c(F)$ under symmetry. Otherwise, it is immediate from A2 that $w_c(ac) > w_c(F)$ for all ζ.

[16] The same critical cost obtains for the unilateral deviation to $\langle\{ab, bc\}\rangle$. In fact, since the two low cost countries are symmetric, their roles can be reversed without loss of generality.

Condition 1A:

$$w_c(ab, ac) < w_c(ab)$$

Note that condition 1A is equivalent to $\zeta^{ab} < \zeta^{ab,ac}$ and if it holds then the hub and poke arrangement $\langle\{ab, ac\}\rangle$ cannot be an equilibrium since country c prefers being a non-member to being a spoke. We can now state:

Proposition 7: *Suppose condition 1A holds. If $\zeta \leq \zeta^{ab}$ then free trade is a stable equilibrium of the FTA game; otherwise the bilateral FTA $\langle\{ab\}\rangle$ is stable.*

Proposition 5 and 7 together imply the following:

Corollary 1: *If condition 1A holds, FTAs act as strong building blocs whenever $\zeta^{\phi} < \zeta < \zeta^{ab}$.*

The above result hinges on the insight that when free trade is not feasible due to the reluctance of the high cost country, the fact that the low cost countries can form a bilateral FTA can make it a willing participant in global free trade since it is worse off as a non-member country than it is under free trade. It is worth noting here that Baldwin (1995) argued that the expansion of a regional trade bloc can induce outsiders to join since their export profits suffer if they stay outside. In our model, a similar logic operates but there are several important differences. First, our logic does not rely on the presence of economies of scale. Second, the formation of FTAs is fully endogenous in our approach. Third, and more importantly, our model highlights the fact that under a purely multilateral approach, a country that is reluctant to liberalize can effectively prevent liberalization between its trading partners and the removal of such 'veto power' can sometimes be necessary to achieve global free trade.

When $\zeta > \zeta^{ab}$ free trade is not feasible under either game. However, the second statement of Proposition 7 implies that there is still a sense in which the freedom to pursue bilateral FTAs proves beneficial. More specifically, the option to form FTAs can lead to welfare-improving trade liberalization that is ruled out by the No FTA game: when $\zeta > \zeta^{ab}$, the FTA game yields $\langle\{ab\}\rangle$ whereas the No FTA game yields $\langle\{\Phi\}\rangle$. Under such circumstances, we say that FTAs act as *partial building blocs* – the option to pursue bilateral FTAs delivers partial trade liberalization whereas the multilateral approach delivers none. The policy implication here is that insisting only on multilateral trade liberalization can sometimes run the risk of ending up with no trade liberalization.

What happens when condition 1A fails? If so, country c prefers $\langle\{ab, ac\}\rangle$ to $\langle\{ab\}\rangle$ and the equilibrium outcome depends on how it ranks $\langle\{ab, ac\}\rangle$ and $\langle\{\Phi\}\rangle$:

Proposition 8: *Suppose condition 1A fails. If $\zeta \leq \zeta^{ab,ac}$ then free trade is a stable equilibrium of the FTA game; otherwise, the hub and spoke arrangement $\langle\{ab, ac\}\rangle$ is stable.*

As is clear, Propositions 7 and 8 are quite similar: the first argues that either free trade or a bilateral FTA between the two low cost countries is stable whereas the second states that either free trade or a hub and spoke arrangement with a low cost country as a hub is stable. Propositions 5 and 8 imply the following:

Corollary 2: *If condition 1A fails, FTAs act as strong building blocs when $\zeta^{\phi} < \zeta < \zeta^{ab,ac}$ and they act as partial building blocs when $\zeta^{ab,ac} < \zeta$.*

However, we should note that when condition 1A fails, FTAs are not always a force for the good. In particular, when $\zeta^{ab,ac} < \zeta < \zeta^{\phi}$ FTAs act as weak stumbling blocs because the hub and spoke arrangement $\langle\{ab, ac\}\rangle$ is stable under the FTA game while free trade is stable under the No FTA game. On other hand, even when condition 1A fails, so long as the high cost country prefers $\langle\{ab, ac\}\rangle$ to $\langle\{\Phi\}\rangle$, there is another sense in which FTAs act as partial building blocs – the FTA game delivers $\langle\{ab, ac\}\rangle$ whereas the No FTA game yields $\langle\{\Phi\}\rangle$.

We next provide a graphical illustration of our main results under asymmetry using linear demand.

6 A linear demand illustration

Suppose $u(x_i) = x_i - \frac{x_i^2}{2} + y$ so that $p_i(x_i) = 1 - x_i$. To ensure that the high cost firm has positive sales in all markets under all trade regimes, we need $\zeta < \overline{\zeta} = \frac{1}{3} - t$: this boundary defines the feasible parameter space in all figures.

We now illustrate stable agreements under the two games of trade liberalization. As is clear from Figure 1, free trade is stable under the No FTA game iff $\zeta \leq \zeta^{\phi} = \frac{1+t}{11}$.

— Insert Figure 1 here –

The region over which free trade is stable under the No FTA game is determined by the high cost country's unilateral incentive to deviate from free trade to no agreement. By contrast, under the FTA game, this region is determined by the high cost country's incentive to deviate from free trade to the bilateral FTA $\langle\{ab\}\rangle$ under which it becomes the excluded country. Figure 2 illustrates the stable agreements under the FTA game. As is clear from Figure 2, free trade is stable under the FTA game iff $\zeta \leq \zeta^{ab} = \frac{3-4t}{17}$. Multiple stable equilibria arise over the darker region: both $\langle\{ab\}\rangle$ and $\langle\{F\}\rangle$ are stable.

— Insert Figure 2 here –

To understand why multiple stable equilibria arise in the darker region, consider possible deviations from free trade in this region. The only deviation from free trade that can potentially occur here is the joint deviation of countries a and c to $\langle\{ac\}\rangle$ and we can show that:

$$w_a(ac) \geq w_a(F) \text{ iff } \zeta \geq \zeta^{ac} = \frac{11t - 2}{10} \tag{31}$$

However, this deviation is *not* self-enforcing since taking country b's announcement as given, country a has an incentive to further deviate from

$\langle\{ac\}\rangle$ to $\langle\{ab, ac\}\rangle$:

$$w_a(ab, ac) - w_a(ac) = \frac{t(6 + 15t - 2\zeta)}{32} > 0 \text{ for all } \zeta \tag{32}$$

Now consider possible deviations from $\langle\{ab\}\rangle$. First, all countries have an incentive to jointly deviate from $\langle\{ab\}\rangle$ to $\langle\{F\}\rangle$. However, as indicated above, this deviation is not self-enforcing when $\zeta \geq \zeta^{ac}$ since countries a and c have an incentive to further deviate from $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$. Second, the joint deviation of countries a and c from $\langle\{ab\}\rangle$ to $\langle\{ab, ac\}\rangle$ is self-enforcing and it determines the left hand-side border of the darker region:

$$w_c(ab, ac) - w_c(ab) \geq 0 \text{ iff } \zeta \leq \zeta^{ab-ab,ac} = \frac{6 - 19t}{34} \tag{33}$$

Figure 1 and Figure 2 together imply that free trade is stable when production technologies are relatively similar across countries. Furthermore, as Figure 2 shows that when country c is sufficiently high cost FTAs act as partial building blocs: multilateral free trade is infeasible and no agreement $\langle\{\Phi\}\rangle$ obtains under the No FTA game whereas the bilateral FTA $\langle\{ab\}\rangle$ emerges under the FTA game. When there exist multiple stable equilibria, there are two distinct scenarios to be considered. In scenario I, we assume that over the darker region in Figure 2, free trade is stable whereas in scenario II the bilateral FTA $\langle\{ab\}\rangle$ is stable. We consider each in turn.

— Insert Figure 3 here –

Under scenario I, Figure 3 shows that free trade is stable over a much *larger* region under the FTA game. Since the high cost country prefers $\langle\{\Phi\}\rangle$ to $\langle\{ab\}\rangle$, its incentive to unilaterally deviate from free trade is greater under the No FTA game relative to the FTA game. As a result, FTAs act as strong building blocs when $\zeta^{\phi} < \zeta < \zeta^{ab}$. It is worth emphasizing that this result obtains when the cost disadvantage of country c is of an intermediate magnitude. If country c is too high cost, it loses too little in export markets from the preferential trade liberalization that occurs between countries a and b and for domestic surplus considerations it prefers to retain its own tariffs. Similarly, when its cost is quite low it prefers to form FTAs with both its trading partners.

Now consider scenario II where $\langle\{ab\}\rangle$ rather than $\langle\{F\}\rangle$ obtains as the stable agreement over the darker region of Figure 2. A comparison of the two trade liberalization games under scenario II is presented in Figure 4.

— Insert Figure 4 here –

Under scenario II, we highlight three different regions in Figure 4. In region A, $\langle\{F\}\rangle$ is a stable agreement of the No FTA game whereas $\langle\{ab\}\rangle$ obtains under the FTA game. As a result, in region A, FTAs act as stumbling blocs for multilateral trade liberalization. However, even under scenario II, FTAs continue to act as building blocs: in region B of Figure 4, they act as strong building blocs whereas in region C they act as partial building blocs (since the bilateral FTA $\langle\{ab\}\rangle$ obtains under the FTA game whereas no agreement $\langle\{\Phi\}\rangle$ obtains under the No FTA game). How does the option to pursue FTAs affect the welfare of individual countries and that of the world as a whole?

Proposition 9: *Under linear demand, the following obtains: (i) when FTAs act as building blocs (strong or partial), the option to form bilateral FTAs benefits low cost countries (as well as the world as a whole) whereas it hurts the high cost country and (ii) when FTAs act as stumbling blocs, all countries lose from being able to form bilateral FTAs.*

Why does the high cost country lose when FTAs act as building blocs? To see why, first note that, for all $t > 0$, the high cost country always prefers no agreement to the bilateral FTA $\langle\{ab\}\rangle$. Thus when FTAs act as partial building blocs, they necessarily make the high cost country worse off. Recall that FTAs act as strong building blocs only when $w_c(\Phi) > w_c(F) > w_c(ab)$. So here too the high cost country loses from the option to form FTAs. Finally, note that when FTAs act as stumbling blocs (as in scenario II), all countries are willing to jointly deviate from $\langle\{ab\}\rangle$ to $\langle\{F\}\rangle$. However, due to the incentives of countries a and c to further deviate from $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$, the bilateral FTA $\langle\{ab\}\rangle$ obtains as a stable equilibrium.

7 Concluding remarks

This paper contributes to the long-standing debate regarding the relationship between preferential and multilateral trade liberalization by analyzing two trade policy games: one where countries can choose between both types of trade liberalization and another where they can only pursue the multilateral route. An important aspect of the paper is that it explicitly models the process of FTA formation and allows each country to form more than a single FTA. Furthermore, to capture FTA formation in a realistic fashion, we focus on Nash equilibria that are immune to self-enforcing coalitional deviations.

At a general level, our analysis sharpens the stumbling versus building bloc debate regarding FTAs by highlighting conditions under which each of the two effects is likely to obtain. More specifically, when countries are relatively symmetric, the option to pursue FTAs hinders the obtainment of global free trade in the sense that free trade is no longer uniquely stable when FTAs are permitted. On the other hand, free trade is harder to sustain under asymmetry and FTAs can actually be desirable from a world welfare perspective. In fact, we show that there indeed exist circumstances where global free trade obtains as an equilibrium *only if* countries are free to form bilateral FTAs. We also find that FTAs can deliver welfare improving trade liberalization when multilateral free trade is infeasible – i.e. it may indeed be better to have some trade liberalization (even though it occurs on a preferential basis) as opposed to none. Another interesting result of the paper is that a hub and spoke type trading regime is not a stable equilibrium under symmetry. Finally, it is worth noting that FTAs *never* act as *strong stumbling blocs* in our model – i.e. it is never the case that the FTA game yields no agreement whereas the No FTA games yields free trade.

To allow for greater richness in the modeling of FTA formation, we have kept the underlying model of intraindustry trade as simple as possible. It is important to determine whether our approach to FTA formation can be fruitfully applied to general equilibrium models of inter and intraindustry trade (such as the Ricardian model of comparative advantage and the monopolistic competition model of trade in differentiated goods). We leave this for future research.

8 Appendix

Trade liberalization increases welfare

Differentiating world welfare with respect to t gives:

$$\frac{dWW}{dt} = \frac{dS_i}{dt} + \sum_{z \neq i} \frac{d\Pi_{zi}}{dt} \tag{34}$$

Using $u' = p_i$ and $x_i = x_{ii} + \sum_{z \neq i} x_{zi}$, we have

$$\frac{dS_i}{dt} = \left[1 - \frac{dp_i}{dt}\right][x_i - x_{ii}] + [p_i - \zeta]\frac{dx_{ii}}{dt} + t\left[\frac{dx_{ji}}{dt} + \frac{dx_{ki}}{dt}\right] \tag{35}$$

Also note that

$$\sum_{z \neq i} \frac{d\Pi_{zi}}{dt} = \left[\frac{dp_i}{dt} - 1\right][x_i - x_{ii}] + [p_i - \zeta - t]\left[\frac{dx_{ji}}{dt} + \frac{dx_{ki}}{dt}\right] \tag{36}$$

where we have made use of the first order conditions for profit maximization for each firm. From equations (34) through (36) we have:

$$\frac{dWW}{dt} = [p_i - \zeta]\frac{dx_{ii}}{dt} + [p_i - \zeta]\sum_{z \neq i} \frac{dx_{zi}}{dt} \tag{37}$$

Using $x_i = x_{ii} + \sum_{z \neq i} x_{zi}$, the following is immediate:

$$\frac{dWW}{dt} = [p_i - \zeta]\frac{dx_i}{dt} < 0 \text{ since } \frac{dx_i}{dt} < 0 \tag{38}$$

i.e. lowering tariff improves world welfare. Analogous arguments establish that (i) the lowering of its tariff on an MFN basis must also improve world welfare and that (ii) it is socially optimal to set an FTA's external tariff to zero.

When countries are asymmetric, trade liberalization increases welfare so long as it increases the aggregate output of good x. Only when liberalization is biased in favor of a high cost country can it be the case that $\frac{dx_i}{dt} > 0$.

Proof of Proposition 6

First consider country a's unilateral deviation from $\langle\{F\}\rangle$ to $\langle\{bc\}\rangle$. The following is immediate from (21):

$$\lim_{\zeta \longrightarrow 0} [w_a(bc) - w_a(F)] < 0 \tag{39}$$

From assumption 2 we know

$$\frac{\partial [w_a(bc) - w_a(F)]}{\partial \zeta} < 0 \tag{40}$$

Inequalities (39) and (40) together imply that

$$w_a(bc) < w_a(F) \text{ for all } \zeta \tag{41}$$

Since $\zeta_a = \zeta_b = 0$, starting at free trade, neither country a nor b has an incentive to unilaterally break any of its FTAs. Recall from (26) that low cost countries have no incentives to unilaterally or jointly deviate from $\langle\{F\}\rangle$ to $\langle\{\Phi\}\rangle$.

We now rule out deviations of country a from $\langle\{F\}\rangle$ to $\langle\{ab, bc\}\rangle$. It is immediate from (22) that

$$\lim_{\zeta \longrightarrow 0} [w_a(ab, bc) - w_a(F)] < 0 \tag{42}$$

From assumption 2 we have

$$\frac{\partial [w_a(ab, bc) - w_a(F)]}{\partial \zeta} < 0 \tag{43}$$

Inequalities (42) and (43) together imply that

$$w_a(ab, bc) < w_a(F) \text{ for all } \zeta \tag{44}$$

As a result, a low cost country has no incentive to unilaterally deviate from $\langle\{F\}\rangle$ to $\langle\{ab, bc\}\rangle$. This also implies that countries a and c have no incentive to jointly deviate from free trade $\langle\{F\}\rangle$ to $\langle\{ab, bc\}\rangle$.

Next, consider country a's deviation from free trade $\langle\{F\}\rangle$ to $\langle\{ac, bc\}\rangle$. We know from (23) that $ww(F) > ww(ac, bc)$. Also, $w_c(F) < w_c(ac, bc)$. This implies that countries a and b must both be worse off under $\langle\{ac, bc\}\rangle$ relative to $\langle\{F\}\rangle$

$$w_a(ac, bc) = w_b(ac, bc) < w_a(F) \tag{45}$$

Therefore, there exist no unilateral or joint deviations from $\langle\{F\}\rangle$ to $\langle\{ac, bc\}\rangle$.

Thus, all feasible coalitional deviations have been ruled out except for the following:

- (JD1): joint deviation of countries a and b from $\langle\{F\}\rangle$ to $\langle\{ab\}\rangle$.
- (JD2): joint deviation of countries a and c from $\langle\{F\}\rangle$ to $\langle\{ac\}\rangle$.

From assumption 3 we know that, even if these two deviations were to occur, taking countries b and c's announcements as fixed, country a has an incentive to further deviate from from $\langle\{ab\}\rangle$ in JD1 (or $\langle\{ac\}\rangle$ in JD2) to $\langle\{ab, ac\}\rangle$. Therefore, neither JD1 nor JD2 is self-enforcing.

Proof of Proposition 7

Proposition 6 states that countries a and b will not deviate from free trade either unilaterally or jointly. To prove that free trade is stable if $\zeta \leq \zeta^{ab}$ we only need to rule out unilateral deviations on the part of country c. First note that if $\zeta \leq \zeta^{ab}$ then country c will not break both FTAs and become a non-member. Second, it has no incentive to break one of its FTAs and become a spoke under a hub and spoke arrangement (Lemma 3).

Now we show that $\langle\{ab\}\rangle$ is stable when $\zeta > \zeta^{ab}$. Since

$$w_a(ab) = w_b(ab) > w_a(\Phi) = w_b(\Phi) \tag{46}$$

countries a and b have no unilateral or coalitional incentive to deviate from $\langle\{ab\}\rangle$ to $\langle\{\Phi\}\rangle$.

We now consider self-enforcing coalitional deviations from $\langle\{ab\}\rangle$. Since condition 1 holds for all ζ, country c has no incentive to jointly deviate with country a (or b) from $\langle\{ab\}\rangle$ to $\langle\{ab, ac\}\rangle$ (or $\langle\{ab, bc\}\rangle$). Next, note that the joint deviation of all countries from $\langle\{ab\}\rangle$ to $\langle\{ac, bc\}\rangle$ is not self-enforcing because countries a and b have a joint incentive to further deviate from $\langle\{ac, bc\}\rangle$ to $\langle\{F\}\rangle$. Furthermore, when $\zeta > \zeta^{ab}$ country c has no incentive to jointly deviate with country a and b from $\langle\{ab\}\rangle$ to $\langle\{F\}\rangle$. Finally, suppose that countries a and c have an incentive to jointly deviate from $\langle\{ab\}\rangle$ to $\langle\{ac\}\rangle$. It is immediate from assumption 3 that country a wants to further deviate from $\langle\{ac\}\rangle$ to $\langle\{ab, ac\}\rangle$. As a result, the initial deviation from $\langle\{ab\}\rangle$ to $\langle\{ac\}\rangle$ is not self-enforcing and $\langle\{ab\}\rangle$ is stable.

Proof of Proposition 8

The first statement of Proposition 8 is immediate from Proposition 6 and the definition of $\zeta^{ab,ac}$. Consider the second part. If $\zeta^{ab,ac} < \zeta$ country c has no incentive to deviate jointly with country b from $\langle\{ab, ac\}\rangle$ to $\langle\{F\}\rangle$. Also, if condition 1A fails then country c has no incentive to deviate (either unilaterally or jointly with a) from $\langle\{ab, ac\}\rangle$ to $\langle\{ab\}\rangle$. Moreover, a simple extension of assumption 3 to asymmetry implies that country a has no incentive to deviate from $\langle\{ab, ac\}\rangle$ to $\langle\{ab\}\rangle$ or $\langle\{ac\}\rangle$:

$$w_a(ab) - w_a(ab, ac) \leq w_a(ac) - w_a(ab, ac) < 0 \text{ for all } \zeta \tag{47}$$

Inequalities (47) and $w_a(ab) > w_a(\Phi)$ imply that country a has no incentive to deviate from $\langle\{ab, ac\}\rangle$ to $\langle\{\Phi\}\rangle$. Now consider the joint deviation of countries b and c from $\langle\{ab, ac\}\rangle$ to $\langle\{\Phi\}\rangle$. Note that even if this deviation occurs, taking a's announcement as given, b has an incentive to deviate further from $\langle\{\Phi\}\rangle$ to $\langle\{ab\}\rangle$. As a result, the initial deviation is not self-enforcing. The only remaining possible deviation is the joint deviation of countries b and c from $\langle\{ab, ac\}\rangle$ to $\langle\{ac, bc\}\rangle$. However, this deviation is not self enforcing since country b has an incentive to further deviate from $\langle\{ac, bc\}\rangle$ to $\langle\{F\}\rangle$. As a result, if $\zeta^{ab,ac} < \zeta$ the pair of bilateral FTAs $\langle\{ab, ac\}\rangle$ is stable.

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] Baldwin, Richard, 1995. "A Domino Theory of Regionalism", in Expanding European Regionalism: The EU's New Members, Baldwin, R, P. Haaparanta and J. Kiander (eds.), Cambridge University Press, 1995.

[4] Bergstrom, Theodore and Hal R. Varian, 1985. "When are Nash Equilibria Independent of the Distribution of Agents' Characteristics." Review of Economic Studies LII, 715-718.

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

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

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

[8] Bond, Eric W., and Constantinos Syropoulos, 1996. "The Size of Trading Blocs: Market Power and World Welfare Effects." Journal of International Economics 40, 411-437.

[9] Brander, J. A., Krugman, P.R., 1983. A 'Reciprocal Dumping' Model of International Trade. Journal of International Economics 15, 313-323.

[10] 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.

[11] Conconi, Paola and Carlo Perroni, 2003. "Self-Enforcing International Agreements and Domestic Policy Credibility," mimeo.

[12] Crawford, Jo-Ann and Roberto V. Fiorentino, 2005. "The Changing Landscape of Regional Trade Agreements," WTO Discussion Paper No. 8.

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

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

[15] Duttagupta, Rupa and Arvind Panagariya, 2006. "Free Trade Areas and Rules of Origin: Economics and Politics," mimeo.

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

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

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

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

[20] Jackson, Matthew and Asher Wolinsky, 1996. "A Strategic Model of Social and Economic Networks." Journal of Economic Theory 71(1), 44-74.

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

[22] 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.

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

[24] Maggi, Giovanni and Andres Rodiguez-Clare, 1998. "The Value of Trade Agreements in the Presence of Political Pressures." Journal of Political Economy 106, 574-601.

[25] Maggi, Giovanni and Andres Rodiguez-Clare, 2007. "A Political Economy Theory of Trade Agreements," American Economic Review 97, 1374-1406.

[26] McLaren, John, 1997. "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade." American Economic Review 87, 400-420.

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

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

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

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

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

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

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

[34] World Trade Organization, 2003. World Trade Report.

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

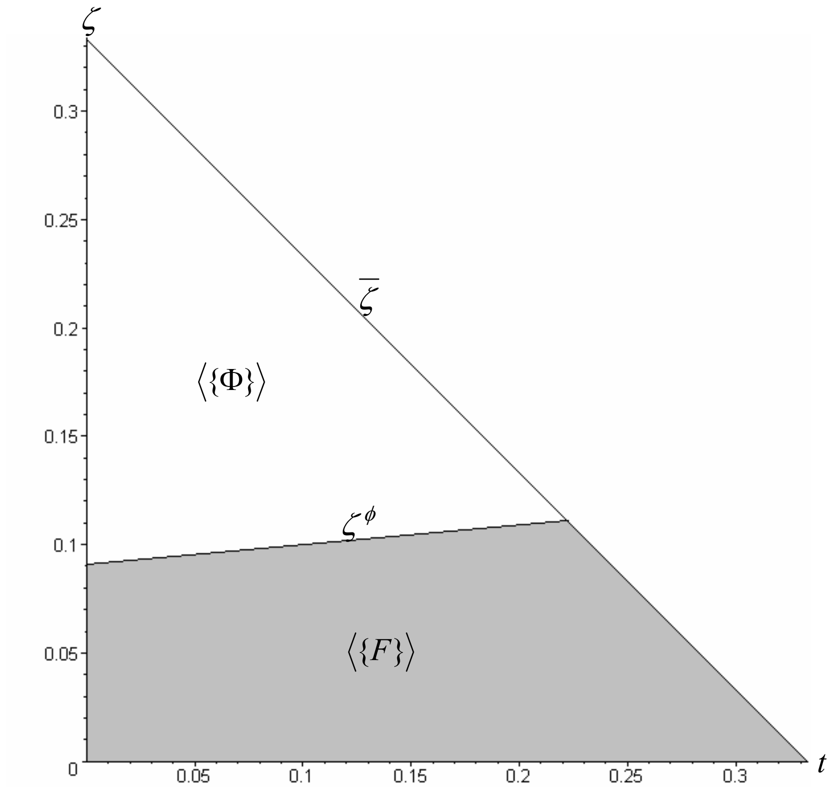

Figure 1: Stable agreements under the No FTA game

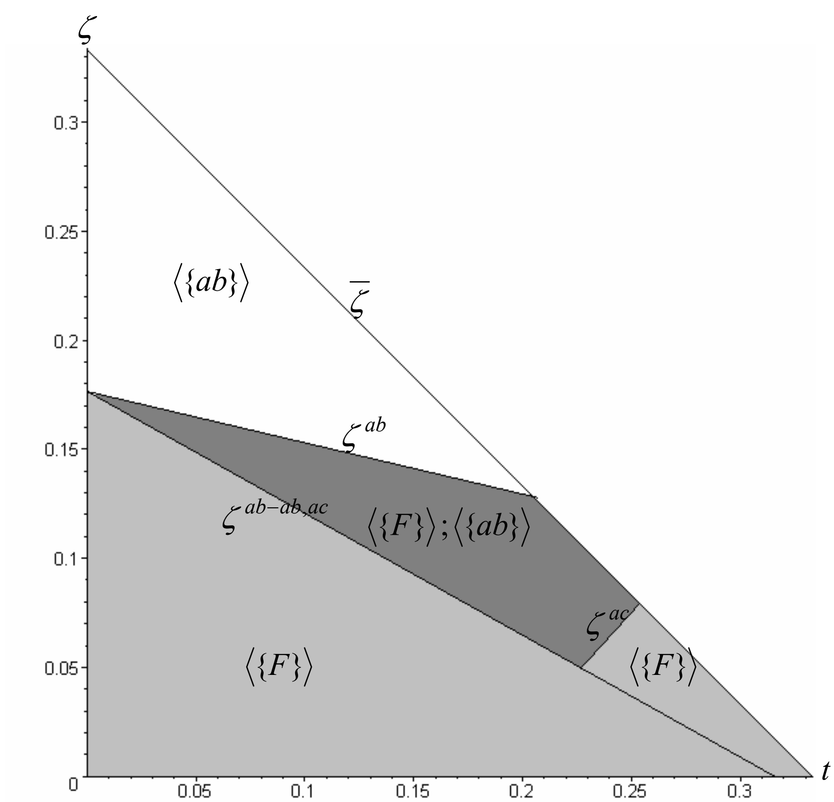

Figure 2: Stable agreements under the FTA game

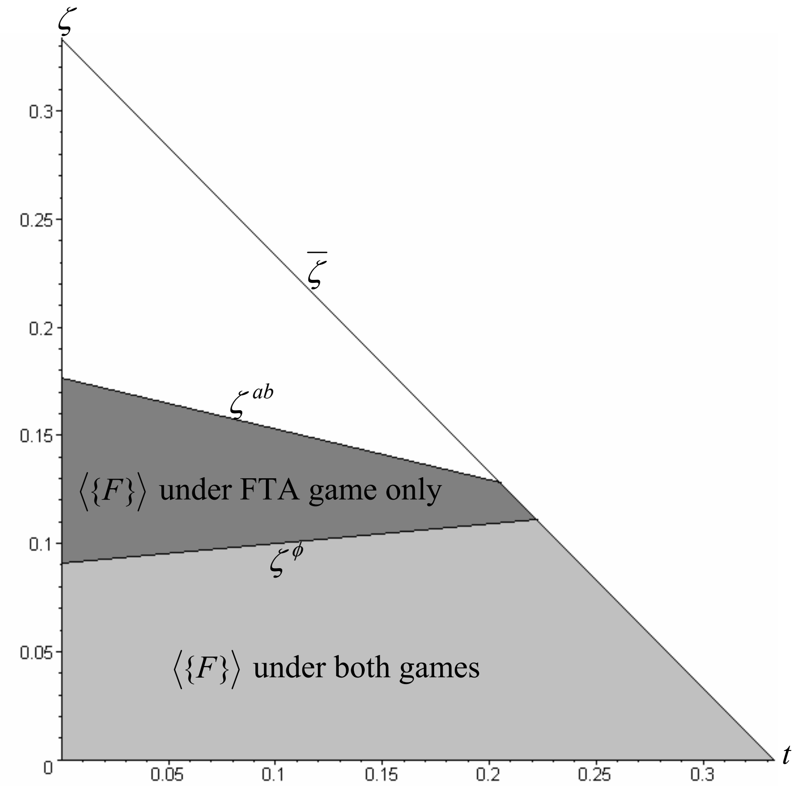

Figure 3: FTAs as Strong Building Blocs

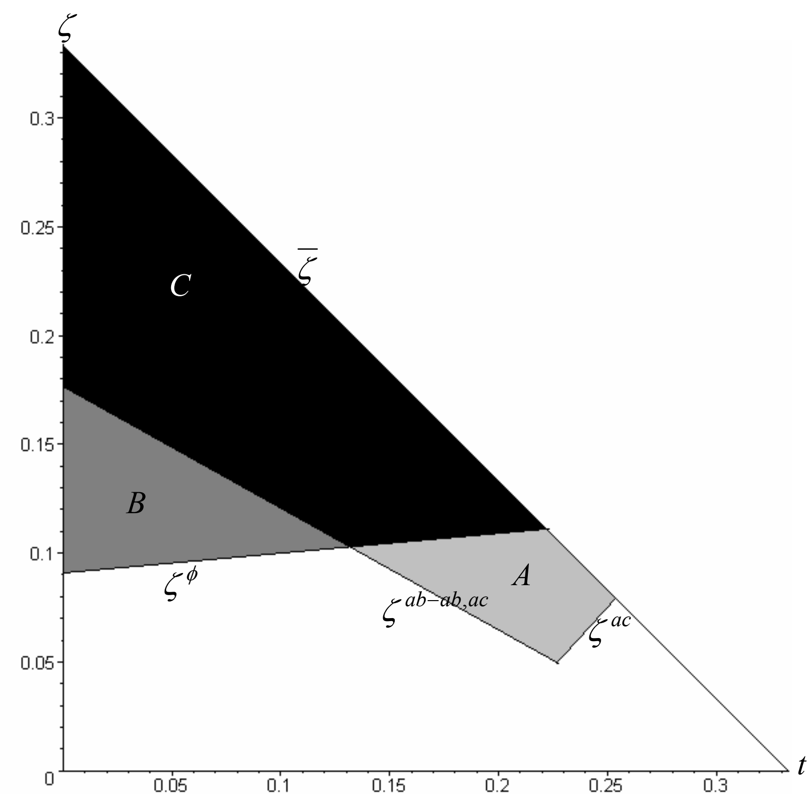

Figure 4: Building and Stumbling Blocs