Why are Trade Agreements Regional?[1]

Ben Zissimos

Vanderbilt University

November 2006

Abstract: This paper shows how distance may be used to coordinate on a unique equilibrium in which trade agreements are regional. Trade agreement formation is modeled as coalition formation. In a standard trade model with no distance between countries, a familiar problem of coordination failure arises giving rise to multiple equilibria; any one of many possible trade agreements can form. With distance between countries, and through strategic interaction in tariff setting, regional trade agreements generate larger rent-shifting effects than non-regional agreements, which countries use to coordinate on a unique equilibrium. Under naive best responses, regional agreements give way to free trade.

Keywords: Coordination, free trade, gradual trade liberalization, preferential trade agreement, regionalism.

JEL Classification Numbers: F02, F13, F15, C73.

[1]I would like to thank Rick Bond, Andrew Daughety, Bill Hutchinson, Sajal Lahiri, Ron Jones, Jennifer Reinganum, Quan Wen and Anne van den Nouweland for useful comments and conversations on this paper. I would also like to thank seminar participants at the International Industrial Organization Conference, Georgia Tech (2004), the Midwest International Economics Meetings at Vanderbilt (2004), the University of Oregon, Eugene, and at Southern Illinois University for their comments.

1 Introduction

Why are trade agreements (TAs) regional?[2] Prominent examples of TAs are the European Union (EU), the Mercado Comun del Cono Sur (MERCOSUR) and the North American Free Trade Agreement (NAFTA). In all three cases, members share common borders. Wider evidence that TAs are regional is provided by WTO (2000), a report titled "Mapping of Regional Trade Agreements", in which each of the 150 agreements notified to the WTO at the time of publication is represented in map form. It shows that the members of continental TAs share common borders, and that when a group of islands forms a TA they tend to be geographically close to each other.[3] It is well recognized that the 'pure' economic gains through trade creation from TAs are likely to be higher within regions than between them; see Wonnacott and Lutz (1987), Krugman (1991), and Summers (1991). And recent econometric work shows that (the inverse of) distance is a good predictor of TA membership; see Baier and Bergstrand (2004). But by what process do countries specifically choose to form regional TAs? This is a question that has not been addressed directly in the past literature.

This paper argues that regional TA formation solves a coordination problem at the heart of the TA formation process. Such agreements have a large number of possible permutations. And, as often acknowledged, countries cannot write binding contracts over tariffs. How, then, does an agreement between a specific set of countries to coordinate over the setting of a specific set of tariffs (both internal and external) come about? This issue has not been addressed in the past literature on TA formation.

There is undoubtedly significant 'pre-play communication' between policy-makers before a TA is formed. This point is used in the past literature to set aside the issue of coordination failure. But by setting the issue aside, the past literature suppresses a potentially significant explanation for why regionalism is part of the TA formation process. In fact, the need for pre-play communication implies that there is a coordination problem to be resolved as part of that process. The main point brought to light in this paper, by setting coordination of an agreement centre stage, is that the geographical organization of countries into regions helps

[2]TA is a 'catch all' term that refers to all agreements in which a group of countries commit to trade among members preferentially. This encompasses free trade agreements (FTAs) in which members agree to remove internatal tariff barriers but set external tariffs independently, and customs unions (CUs) which are like FTAs with the additional requirement that members coordinate on common external tariffs. In practice, FTAs are more common but most of the academic literature focuses on CUs because they are analytically easier to handle. To focus the discussion on the regional nature of these agreements rather than the technical details of their operation, we will use the catch-all term TA wherever possible.

[3]The Caribbean Common Market (CARICOM) is an example of a group of close islands that have a TA.

to determine which other countries a given country chooses to approach when deciding to form a TA.[4]

The model is based on Brander and Spencer (1984) and Yi (1996). Brander and Spencer (1984) show, in a two country model of cross-hauling, that rents made by foreign firms in the domestic market can be shifted back home by the government using tariffs. Yi (1996) uses a Brander-Spencer type model to show that a group of countries may obtain a higher payoff from TA formation than from moving to free trade. The present paper takes a special case of Yi's model and extends it by putting it in a regional setting.[5]

A new effect is revealed in this present paper by putting the Brander-Spencer-Yi model in a regional setting. One of Yi's key results (his Proposition 8) shows that a country would always prefer to leave its own TA in order to join another TA of equal or larger size. While fully acknowledging its importance we will show, in Proposition 5 of this present paper, that Yi's result is overturned in a regional setting. Because the transport costs of inter-regional trade are higher than for intra-regional trade, there is more scope for rent-shifting within a region than across regions (since more rents are dissipated during inter-regional transportation). Moreover, because TA formation eliminates harmful rent shifting between members, and has beneficial terms-of-trade effects against non-members, the benefit to a member of joining a regional TA of a given size is greater than the value to a TA across regions. This effect tends to push the countries of a region towards the formation of a regional TA.[6]

To get a better understanding of the intuition behind this result, consider the original proposals for NAFTA - the North *Atlantic* Free Trade Agreement - between Canada, the UK and the US. Interpreted within the context of the present model, Canada and the US would have been better off if the UK had formed a TA with them, but the UK obtained a

[4]There is a literature that looks at the feasibility of multilateral trade agreements when countries cannot write binding contracts over tariffs; see for example Bagwell and Staiger, (1997a,b), (1999), Bond and Syropoulos (1996), Bond (2001), Bond, Syropoulos and Winters (2001) and Bond, Riezman and Syropoulos (2004). All of these previous papers look at how agreements between sufficiently patient countries may be sustained through repeated interactions in the face of a short run incentive to deviate. In the model of this present paper, there is no short-run incentive to deviate. The problem focused on in this paper is not whether the agreement can be made self enforcing but whether a country is able to form an agreement with the other countries that it would like to have as members - the problem of coordination failure.

[5]Yi's (1996) focus is on how 'open regionalism' can help with the attainment of free trade compared to the outcome when TAs are exclusive. While the present paper draws on Yi's analysis of exclusive regionalism, it does not address the question of whether open regionalism would be benefical in a regional setting.

[6]At a more abstract level, Yi's result breaks down because of a shift away from the assumption that all players are ex ante homogeneous (where players are obviously national governments here) to the present setting where regionalism introduces a degree of heterogeneity between players.

higher payoff from the formation of an agreement with nearby EU nations because the gains to elimination of rent shifting within Europe were of greater value to the UK.[7]

The problem of coordination has attracted significant attention in the abstract literature on coalition formation. In this theoretical literature it is widely understood that many equilibria can arise through the inability to coordinate or commit (see Bloch 2003 and Yi 2003 for reviews). The problem is captured formally by the requirement that each player simultaneously and noncooperatively writes down a list of other players with whom he would like to form a coalition. In this setting, any one of many possible coalition configurations may arise in equilibrium.

To introduce the problem of coordination failure in the present context, TA formation is modelled based on Hart and Kurz's (1983) *simultaneous move exclusive membership game*. In their original game, without communicating, each player writes down a list of other players with whom he would like to form a coalition. The lists form intersecting sets of players and each of the intersecting sets forms a coalition. But if two players fail to name each other then neither ends up in the same coalition even if it would be mutually beneficial. In the model of this present paper, each country writes down a list of others with whom it would like to form a TA. When there are no transport costs, countries are ex ante symmetrical and the problem of coordination failure may arise between them. On the other hand, when transport costs of trading between regions are sufficiently large (but not large enough to prohibit trade between regions) countries use the difference in rent-shifting effects within and between regions to coordinate on regional TA formation; TAs form simultaneously, one in each region, and each TA includes all countries in that region. This is the sense in which coordination failure is resolved when transport costs are introduced to the model.[8]

A key concern about the implications of regionalism is whether or not it is consistent with the gradual attainment of efficiency; whether trade blocks are 'stepping blocks or stumbling blocks in the path to free trade' (Bhagwati 1992). The insights gained about regional TA formation from our present model are helpful to this debate. An implication of Yi's Proposi-

[7]The underlying intuition is robust to the fact that the NAFTA proposals were obviously for an FTA while the EU is a CU. In a broader setting, the choice of trading arrangement may have a significant bearing on the outcome, as shown by Riezman (1999) who endogenizes the decision by countries over whether to adopt a CU or FTA, showing that this may affect whether free trade can be reached. (See also Bloch's 2003 discussion of CUs versus FTAs, and Bond, Riezman and Syropoulos 2004.)

[8]It would be nice if we could have an intermediate step in the analysis wherein the regional dimension is introduced to the model but the game of coalition formation is the same as used by Yi (1996), namely, Bloch's (1996) size announcement game. Unfortunately, this is not possible since Bloch's size announcement game does not admit an equilibrium once transport costs are introduced to the model.

tion 8 is that an equilibrium TA structure must be asymmetric. Countries use the advantage in the sequence of TA formation that they are exogenously granted to form a larger TA. The countries in the larger TA are better off even than under free trade because they enjoy more favorable terms-of-trade effects over non-member countries. As a result, trade blocks are stumbling blocks in the path to free trade. In the present paper no such advantages arise due to the fact that TA formation is simultaneous and because each country is uncertain about TA formation beyond its own. As a result TA formation can be symmetric, with no larger TA arising that would prefer the status quo to free trade. In that case regional trade blocks do ultimately facilitate free trade.[9]

It is important to emphasize that because the model presented in this paper is based on an example the results are at best suggestive, and may not be robust to certain alternative specifications. Nevertheless, even though the policy analysis is based around intra-industry trade in a single sector, and strong assumptions are made about functional forms, the results seem intuitively plausible, and may be indicative of a driving force towards regionalism for which there appears to be statistical evidence.

The paper proceeds as follows. Section 2 introduces the basic trade model. Section 3 then explores the economic implications of TA formation in the model, and results are established which are useful in the characterization of equilibrium TA formation. Section 4 introduces the TA formation game. Based on the TA formation game, Section 5 then shows that when all countries are symmetrical and there are no transport costs (i.e. zero distance) between them, there are multiple equilibria and no predictions can be made as to which will prevail. Section 6 shows that when transport costs are introduced, this provides a mechanism through which countries are able to coordinate on a unique equilibrium, inducing regional TAs to form. Section 7 then examines the extent to which regional TA formation may subsequently give way to free trade. Conclusions are drawn in Section 8.

[9]This effect is closely related to an effect demonstrated by Riezman (1999) in a different setting. The naive best response dynamics are adapted from Bala and Goyal's (2003) model of network formation.

2 The Basic Model

We will work with a familiar model of international trade based on Cournot competition.[10] Each country has a representative consumer, firm, and government. Let N be the set of countries. Each type of agent, consumer, firm and government, is denoted by its corresponding country identifier as $i \in N$. Firms produce a homogenous product.

There are just six countries; $N = \{1, 2, 3, 4, 5, 6\}$. The *regional structure* partitions N into regions such that $R_1 = \{1, 2, 3\}$ and $R_2 = \{4, 5, 6\}$. Regions are some distance apart from one another. Let d_{ij} measure the distance between any two countries $i, j \in N$. If countries i and j are *not* in the same region let $d_{ij} = d$, while if i and j are in the same region let $d_{ij} = 0$.[11]

The model of TA formation is dynamic with three periods. Within a period, the sequence of events is as follows. First, TA formation takes place. Next, taking TAs as given, firms make production decisions. Finally, consumption takes place.

TA formation, modelled as a dynamic game of coalition formation, will be described in detail in Section 4. Before that, we will set out the model of production, taking tariffs as given, then use this to determine tariffs. The model of production and trade subject to tariffs will then provide a concrete source of gains on which to base the payoffs in the game of TA formation.

2.1 Preferences

There are two goods in the model, denoted M and X. Good M is chosen as the numeraire. Countries are endowed with equal quantities of M, the role of which is to ensure that trade accounts balance in equilibrium. The term M_i measures consumption in country i of M.

[10] Use of the Cournot oligopoly model as a general trade model has been criticized on empirical grounds because a relatively small share of overall economic activity has an oligopolistic structure. But it is recognized that results obtained using this model are indicative of those found using a more general class of models. It is used in this present setting for tractability, because it yields an expression for the tariff coordinated by a CU that depends only on the number of its own members. This feature is not available in any of the other commonly used models of CU formation.

[11] When countries are ex ante homogeneous (no transport costs) it is possible to analyze a model of N countries in an oligopoly trade model, as in Yi (1996) and Goyal and Joshi (2005). When countries exa ante heterogeneous (positive transport costs between regions) the model becomes more complicated particularly in terms of tariff setting. Bond (2001) considers regionalism versus multilateralism in a model of four countries, with two in each region. I too focus on two geographically distinct regions, but I use six countries in order to examine trade block formation within a region that does not encompass the whole region. For this, it is necessary to have (at least) three countries in each of the two regions.

All the firms in the model, one in each country, produce the homogeneous product X. The term X_i measures consumption of X in country i. This quantity is given by the summation

$$X_i = \sum_{j \in N} x_{ij}, \tag{1}$$

where x_{ij} is the quantity produced by the firm in country j for the market in country i.

Preferences are approximated by the following quasi-linear function:

$$u_i = v(X_i) + M_i = eX_i - \frac{1}{2}X_i^2 + M_i. \tag{2}$$

where e is a parameter. This functional form is relatively simple, focusing attention on the impact of product differentiation by distance.[12]

The inverse demand curve of consumer i is obtained in the usual way by differentiating (2) with respect to x_{ij}, assuming that firms can segment markets and all other outputs remain constant:

$$p_i\left(X_i\right) = \frac{dv}{dx_{ij}} = e - X_i. \tag{3}$$

2.2 Production with Costs as a Function of Tariffs and Distance

Firm j's (marginal) cost to produce a unit of X for sale in country i consists of three components: a private per unit cost, c, which is uniform for all firms; the tariff levied by government i on imports from j, t_{ij}; the transport cost of shipping from j to i, captured simply by d_{ij}. Thus, firm j's per-unit production cost for each market i is given by the function

$$c_{ij} = c + t_{ij} + d_{ij}. \tag{4}$$

Given that firms are able to segment markets, firm j chooses the quantity to produce (and export if $j \neq i$) in order to maximize profits in each market:

$$\text{Max}_{x_{ij}}\ \pi_{ij} = \left(p_i - c_{ij}\right) x_{ij}, \tag{5}$$

where p_i is determined according to the inverse demand curve $p_i\left(X_i\right)$ given by (3).

[12] Yi (1996) has a more general form of this production function which allows X to be horizontally differentiated. The model of this present paper could be extended in that direction but this would complicate the analysis considerably and would risk obscuring the effects resulting from the organization of countries into regions.

Output under Cournot competition is solved for in the usual way. Setting the first-order-condition of (5) equal to zero, the profit maximizing solution for output of firm j in country i, x_{ij}, depends on the output of all the other firms that produce for country i.[13] Solving simultaneously, we obtain the following solution for x_{ij}:

$$x_{ij} = \frac{(e-c) + \sum_{k \in N} d_{ik} + \sum_{k \in N} t_{ik}}{n+1} - d_{ij} - t_{ij}. \tag{6}$$

Output by firm j for market i depends negatively on d_{ij} and t_{ij}; the smaller the distance to market, and the lower the tariff, the larger the rents available from shipping to country i and so the higher the quantity produced. In contrast, output by firm j depends positively on the distance from country i to all other markets and the tariff set by country i on imports from all countries other than j. Note that the strength of demand relative to cost helps to determine the rents available to firm j as well; $e - c$ is common to all markets in this model and can be made large enough to ensure that $x_{ij} > 0$ for all i, j.

2.3 Welfare

Profits of domestic firms and tariff revenues are rebated back to consumers. Also, there is perfect competition in the world market for transportation. Thus, country i's welfare, w_i, can be expressed in terms of three economic components: domestic consumer surplus, C_i; the domestic firm's profit at home and abroad, π_{ii} and $\sum_{k \in N \setminus i} \pi_{ki}$ respectively $(k \neq i)$); tariff revenue, T_i:

$$w_i = C_i + \pi_{ii} + \sum_{k \in N \setminus i} \pi_{ki} + T_i + D_i, \tag{7}$$

where $C_i = \frac{1}{2}(e - p_i) X_i$, $T_i = \sum_{k \in N} t_{ik} x_{ik}$. This way of expressing national welfare in terms of these three economic components is familiar from the earlier literature. Because the transport sector is perfectly competitive, goods are delivered at cost and there is no surplus associated with that sector; $D_i = 0$.

3 Optimal Tariffs and Production in the 2-Region Model

At this stage, the structure of TAs in the world economy is taken as given. If countries i and j are both in the same trade agreement then they set $t_{ij} = t_{ji} = 0$. Members of a TA coordinate in setting tariffs on imports from non-members.

[13] This is firm j's reaction function for market i.

The structure of TAs in the world economy is defined as follows. A *TA structure* $B = (B_1, B_2, ..., B_m)$ is a partition of the set of countries N. $B_i \cap B_j = \emptyset$ for $i \neq j$, and $\cup_{i=1}^{m} B_i = N$. A *singleton* is a country that does not coordinate trade policy with others and simply optimizes tariffs on a unilateral basis; a singleton is equivalent to a single-member-country TA. Note that in the present model there can be at most six TAs where each TA is a singleton.

Recall that the location of each country is fixed either in R_1 or in R_2. Therefore, $(B_k \cap R_1) \cup (B_k \cap R_2) = B_k$. It will be helpful in what follows to have notation for the number of TA members that are in a country's own region and the number in the 'other' region. Formally, if $i \in B_k$ and $i \in R_l$ then let b_{ir} be the cardinality of the set $B_k \cap R_l$ and let b_{inr} be the cardinality of the set $B_k \cap R_m$, $l \neq m$. Less formally, b_{ir} measures the number of countries in its own region with which country i shares a TA (including itself) and b_{inr} measures the number of countries in the other region and in i's TA. In the present simple regional set-up, $b_{ir} \in \{1, 2, 3\}$ and $b_{inr} \in \{0, 1, 2, 3\}$. We will refer to a pair $\{b_{ir}, b_{inr}\}$ in which $b_{ir} \in \{1, 2, 3\}$ and $b_{inr} \in \{0, 1, 2, 3\}$ as a *feasible pair*.

3.1 Optimal tariffs

We now determine optimal tariffs; let r stand for regional and nr stands for non-regional. Then t_{ir} is the tariff that country i sets on imports from non-members in the same region and t_{inr} is the tariff set on imports from non-members outside the region. The optimal levels for these tariffs are derived in the next result. First, the following definition will be helpful with writing down the optimal tariff:

$$\Delta(b_{ir}, b_{inr}) \equiv 7 + (1 + (b_{ir} + b_{nr}))(3 + 2(b_{ir} + b_{inr})),$$

where $n + 1 = 7$.

Proposition 1. *Assume that country i belongs to a TA of b_{ir} regional members and b_{inr} non-regional members. Country i's unique optimal external tariff on imports from a non-member in the same region as country i is*

$$\begin{aligned} t_{ir}^{*}\left(b_{ir}, b_{inr}; d\right) &= \frac{\left(1+2\left(b_{ir}+b_{inr}\right)\right)(e-c)}{\Delta\left(b_{ir}, b_{inr}\right)} \\ &+\frac{3+6 b_{ir}+b_{inr}\left(2\left(b_{ir}+b_{inr}\right)-7\right)}{2 \Delta\left(b_{ir}, b_{inr}\right)} d \end{aligned}$$

The unique optimal external tariff imposed by country i on non-members who are not in the same region as country i is

$$\begin{aligned} t_{inr}^{*}\left(b_{ir}, b_{inr}; d\right) &= \frac{\left(1+2\left(b_{ir}+b_{inr}\right)\right)(e-c)}{\Delta\left(b_{ir}, b_{inr}\right)} \\ &-\frac{5+b_{ir}\left(2 b_{ir}-3\right)+2 b_{inr}\left(5+b_{ir}\right)}{2 \Delta\left(b_{ir}, b_{inr}\right)} d . \end{aligned}$$

The most important thing to notice about $t_{ir}^{*}\left(b_{ir}, b_{inr}; d\right)$ and $t_{inr}^{*}\left(b_{ir}, b_{inr}; d\right)$ is that the difference between them depends on d. That is, $t_{ir}^{*}\left(b_{ir}, b_{inr}; d\right)=t_{inr}^{*}\left(b_{ir}, b_{inr}; d\right)$ for $d=0$ and $t_{ir}^{*}\left(b_{ir}, b_{inr}; d\right)-t_{inr}^{*}\left(b_{ir}, b_{inr}; d\right)$ is increasing in $d>0$. Also notice that if $d=0$ then $t_{ir}^{*}=t_{inr}^{*}$ corresponds exactly to the optimal tariff found in previous literature.[14]

3.2 Demand functions by region and TA membership

Let m stand for TA member; let nm stand for non-member. Based on (6), we can now write down expressions for equilibrium output produced by country j for country i along two dimensions; whether or not country j is a member of country i's TA and whether or not country j is in the same region as country i. These expressions are obtained from (6) by appropriate substitution of distance $d_{ij}=0$ or $d_{ij}=d$ and optimal tariffs $t_{ir}^{*}\left(b_{ir}, b_{inr}\right)$ and $t_{inr}^{*}\left(b_{ir}, b_{inr}\right)$.[15]

[14] In the model of Yi (1996), setting $\gamma=1$ (Yi's notation) causes the utility function that he uses to replicate (2), the expression for u_i in the present paper. Then setting $n=6$ the expression for the optimal tariff that Yi presents in *his* Proposition 1 is given by $\tau(k)=(1+2k) / \left(8+3k+2k^{2}\right)$, where $\tau(k)$ is Yi's notation for the optimal tariff and k is the number of countries in the trade block. In the model of the present paper, if we let $d=0$ and $k=b_{ir}+b_{inr}$ then $t_{ir}^{*}=t_{inr}^{*}=(1+2k)(e-c) / \left(8+3k+2k^{2}\right)$. Note that Yi assumes $e-c=1$ (in my notation).

[15] The intermediate step of solving for equilibrium output based on the model's regional structure but for arbitrary tariffs is presented in the appendix.

Write x_{irm} for output produced for country i by a country that is in the same region as country i *and* is a member of country i's TA:

$$x_{irm}\left(b_{ir}, b_{inr}\right) = \frac{2\left(1+b_{ir}+b_{inr}\right)(e-c)+\left(3\left(1+b_{ir}\right)+2b_{inr}\left(\left(b_{ir}+b_{inr}-1\right)\right)\right)d}{\Delta\left(b_{ir}, b_{inr}\right)}. \quad (8)$$

Write x_{irnm} for output produced for country i by a country that is in the same region but *not* a member of country i's TA:

$$x_{irnm}\left(b_{ir}, b_{inr}\right) = \frac{2(e-c)+\left(3+b_{inr}\left(3+2\left(b_{ir}+b_{inr}\right)\right)\right)d}{2\Delta\left(b_{ir}, b_{inr}\right)}. \quad (9)$$

Write x_{inrm} for output produced for country i by a country not in the same region but which is a member of country i's TA:

$$x_{inrm}\left(b_{ir}, b_{inr}\right) = \frac{2\left(1+b_{ir}+b_{inr}\right)(e-c)-\left(5+2b_{ir}^{2}+b_{inr}\left(5+2b_{ir}\right)\right)d}{\Delta\left(b_{ir}, b_{inr}\right)}. \quad (10)$$

Finally, write x_{inrnm} for output produced for country i by a country that is not in the same region and is not a of country i's TA:

$$x_{inrnm}\left(b_{ir}, b_{inr}\right) = \frac{2(e-c)-\left(5+b_{ir}\left(3+2\left(b_{ir}+b_{inr}\right)\right)\right)d}{2\Delta\left(b_{ir}, b_{inr}\right)}. \quad (11)$$

From these expressions we can work out total output as

$$X_i = b_{ir}x_{irm}+\left(3-b_{ir}\right)x_{irnm}+b_{inr}x_{inrm}+(3-b_{inr})x_{inrnm}.$$

3.3 TA Expansion and Welfare

In this subsection I examine the effect on welfare of TA formation and expansion. I restrict attention to a situation where firms in all countries produce positive quantities for all markets. Output levels and hence trade flows are positive even between countries that are in different regions and not members of the same TA.

It can be seen by inspection that trade flows are lowest between countries that are not members of the same TA and are not in the same region; $x_{inrnm}\left(b_{ir}, b_{inr}\right)$ is the smallest of the quantities given by (8)-(11). Also, by (11), $x_{inrnm}\left(b_{ir}, b_{inr}\right)$ is decreasing in d. It follows that placing an upper bound on d ensures that $x_{inrnm}\left(b_{ir}, b_{inr}\right) > 0$ and in turn all trade flows are positive. The next result identifies the upper bound on d.[16]

[16]The reason for restricting attention to $b_{ir} \in \{1,2,3\}$, $b_{inr} \in \{0,1,2\}$ is because there are no non-regional non-members under free trade ($b_{ir} = 3$, $b_{inr} = 3$), and so it does not make sense to calculate a quantity for $x_{inrnm}(3,3)$.

Lemma 2. *Fix* $a > c$. *If* $d \in (0, (e-c)/22)$ *then* $x_{irm}(b_{ir}, b_{inr}) > x_{inrm}(b_{ir}, b_{inr}) > x_{irnm}(b_{ir}, b_{inr}) > x_{inrnm}(b_{ir}, b_{inr}) > 0$ *for* $b_{ir} + b_{inr} \leq 5$, *where* $b_{ir} \in \{1, 2, 3\}$, $b_{inr} \in \{0, 1, 2\}$. *If* $d \in (0, (e-c)/22)$ *then* $t^*_{ir}(b_{ir}, b_{inr}) > t^*_{inr}(b_{ir}, b_{inr}) > 0$.

To restrict attention to positive output levels, the following standing assumption will be imposed throughout.

Assumption 1. $d \in [0, (e-c)/22)$.

In addition, Assumption 1 ensures that tariffs are positive where they are imposed. Thus, TA formation always entails the removal of positive tariffs.

We now turn to look at the effect of TA formation on member and non-member welfare. I follow Yi (1996) by looking first at the effect of TA formation on non-member countries. Recall that in Yi's model, all countries are ex ante homogeneous (no transport costs). Yi shows (in his Proposition 3) that if a TA forms or expands, then non-member countries are adversely affected. I now show that Yi's result extends to the present environment for heterogeneous countries (there are transport costs to trading between countries in different regions).

In the present model TA expansion may occur within a region (in which case b_{ir} increases) or across regions (in which case b_{inr} increases). Thus, I define *TA expansion* entirely in terms of an increase in b_{ir} and b_{inr}; for a TA to expand, b_{ir} and/or b_{inr} must increase.[17] *TA formation* is just a special case of TA expansion in which all members of the TA that forms start as singletons.

Also note that TA expansion only affects non-members through the demand for exports. This is because optimal tariff setting of non-members is unaffected by TA formation. Thus we can evaluate the effect of TA formation on non-members entirely in terms of the effect on non-member exports to the TA, x_{irnm} and x_{inrnm} and hence export profits. The next result shows that both x_{irnm} and x_{inrnm} are globally decreasing in b_{ir} and b_{inr}.

[17] Say that a CU initially has two members, one in each region. Then say that one of the original members of the CU breaks up with its partner in the other region and instead forms a CU with two countries from its own region. Although a new larger CU is created, this is not allowable under the definition of CU expansion since it involves cessation of membership of one of the initial members. This definition is not restrictive in the game of CU formation that I consider below, because once a CU is formed cessation of membership is not allowed.

Proposition 2. *For all feasible pairs* $\{b_{ir}, b_{inr}\}$, $dx_{irnm}/db_{ir} < 0$, $dx_{inrnm}/db_{ir} < 0$, $dx_{irnm}/db_{inr} < 0$, $dx_{inrnm}/db_{inr} < 0$. *A non-member country's volume of exports and export profits to a TA of size* b_{ir}, b_{inr} *is decreasing in* b_{ir} *and decreasing in* b_{inr}. *The expansion or formation of a TA reduces the welfare of non-member countries.*

I now examine the effect of TA formation on the welfare of members. Yi is able to show for his model that the *joint* welfare of countries involved in TA expansion increases (the welfare of existing members and new members). And more generally, if several TAs merge to form a larger TA the aggregate welfare of the member countries increases. Yi remarks that consumer surplus displays a non-monotonicity that is present in underlying optimal common external tariffs; it may first decrease and then increases as its TA expands. A country's export profits, on the other hand, may initially increase but ultimately decreases as the union expands. The present model introduces a further ambiguity because there are two common external tariffs; the one levied on countries in the same region and the one levied on countries in the other region.

Even though the economic environment is made more complicated by the regional dimension of the model, the next result shows that Yi's Proposition 4 extends to the present setting as well.

Proposition 3. *The expansion or formation of a TA increases the aggregate welfare of member countries.*

Despite the levying of different common external tariffs across regions, the same logic that underpinned Yi's Proposition 4 may be applied here too. If a set of countries abolishes tariffs internally and sets common external tariffs to maximize aggregate welfare then their joint welfare must improve. Proposition 3 shows that forming a TA improves joint welfare of member countries even if non-negative tariffs on imports are the only policy tools and even though members and non-members may be in different regions.

So far, we have seen that Yi's results concerning TA expansion in an environment where all countries are ex ante homogeneous carry over to the present setting where countries are ex ante heterogeneous. When a TA expands, this increases the aggregate welfare of the countries in the TA and harms countries that are not members of the TA. Just as in the world where countries are ex ante homogeneous, this implies that the effect of TA expansion on global welfare is ambiguous. The single case in which this ambiguity disappears is the case where TA expansion goes all the way to the grand coalition, which is equivalent to world free trade. Thus, Yi's Proposition 5 carries over to the present setting and is reproduced here for completeness

Proposition 4. *The effects on global welfare of the formation or expansion of TAs are ambiguous, except when the grand TA is created. World welfare is higher under the grand TA (world free trade) than under any other TA structure.*

All of Yi's results that we have examined so far carry over to the present setting. These results have focused on the welfare effects of TA expansion on non-members and on the aggregate welfare of members.

I now focus explicitly on the welfare of individual member countries in the TA formation process. In doing so, I show that a key property of Yi's homogeneous-country model *fails* to hold when transport costs are sufficiently large but still in the range defined by Assumption 1. Of course, Yi's result continues to hold when transport costs are sufficiently small.

Proposition 5. *There exists a unique value $d' \in (0, (e-c)/22)$ such that for $d \in [0, d')$, a country is better off in a (4-country) TA consisting of itself and all 3 countries* in the other region *than in a (3-country) regional TA* in its own region. *For $d \in (d', (e-c)/22)$, a country is better off in a (3-country) regional TA in its own region than in a (4-country) TA consisting of itself and all 3 countries in the other region.*

For $d \in [0, d')$ this result is consistent with Yi's Proposition 8, which says that a member of a TA becomes better off if it leaves its TA to join another TA of equal or larger size. But for $d \in (d', (e-c)/22)$, the result says that a country is better off remaining in a 3-country TA within its own region than it would be if it left its regional TA to form a 4-country TA with all three countries in the other region.

To understand the intuition behind this result, let us consider a member of a regional TA (in its own region), and ask whether it could gain by joining a regional TA in the other region. Say that Country 1 is initially in a TA; $1 \in B_1 = R_1$, and that the countries in the other region form another regional TA, $B_2 = R_2$. Country 1 considers whether it could gain by leaving B_1 and joining B_2. Decompose the process into three steps: (i) Existing members of B_2 abolish tariffs on imports from Country 1 and change tariffs on the other countries in R_1 from $t^*_{inr}(3,0)$ to $t^*_{inr}(3,1)$; (ii) Country 1 abolishes tariffs on all countries in B_2, and levies tariffs at $t^*_{ir}(1,3)$ on its two former TA partners in B_1; (iii) The remaining two members of B_1 change tariffs on the (original) members of B_2, who are located in R_2, from $t^*_{inr}(3,0)$ to $t^*_{inr}(2,0)$ and levy a tariff $t^*_{ir}(2,0)$ on Country 1.

Consider the effect of each of these steps on the welfare of Country 1 for $d \in [0, d')$ and $d \in (d', (e-c)/22]$ respectively. Take $d \in [0, d')$ first. The abolition of tariffs by the members of B_2 has a positive impact on the welfare of Country 1, because Country 1 enjoys greater openness in three markets. Country 1's abolition of tariffs on all three countries in B_2 also improves welfare but the implementation of tariffs on its two former TA partners in B_1 reduces welfare; the net effect is positive because access is increased to three markets while it is reduced in two. Finally, the implementation of tariffs by its two former TA partners in B_1 reduces export profits and hence welfare in Country 1. But the effect on exports of access to its three new partners in B_2 (in step (i)) more than compensates. The positive effect on consumer surplus from net tariff removal in moving to the larger TA is greater than the negative effect on tariff revenue and the loss of domestic profits from greater competition in the domestic market.

Now take $d \in (d', (e-c)/22]$. The impact on welfare for Country 1 of moving from B_1 to B_2 is reversed. As before, the removal of tariffs by Country 1's three new partners in B_2 has a positive impact on export profits while the implementation of tariffs by Country 1's two former partners in B_1 has a negative effect on export profits. But in the presence of transport costs, the implementation of tariffs by two nearby partners has a larger negative effect on export profits than the removal of tariffs by the three new distant partners in the other region. The effect on consumer surplus is still positive but smaller than before and is not large enough to compensate for the loss of export profits.

Thus, a key result of Yi's is overturned in the present model with the introduction of transport costs. This is significant because it shows that a country will not leave a TA in its own region to form or join a TA in the other region, even if the new TA that forms is larger. In Yi's characterization of an equilibrium TA structure, the first TA to form is the largest. This result calls into question whether a country would always agree to join a larger TA.

A natural question to ask next is whether the members of a regional TA would invite a country from the other region to join them. The next result shows that, once again, the answer depends on the size of transport costs.

Proposition 6. *There exists a value* $d'' \in (d', (e-c)/22)$ *such that for* $d \in [0, d'')$ *the highest feasible level of welfare is achieved when a country is a member of a TA with all of its regional partners and one country from the other region while non-members are singletons. For* $d \in (d'', (e-c)/22)$*, the highest feasible level of welfare is achieved when a country is a member of a regional TA (with all members from its own regional and no members from the other region) while non-members are singletons.*

This result is again in keeping with Yi (1996). A group of countries can obtain a higher level of welfare than under free trade by forming a TA while non-members remain as singletons. In the homogeneous-country model, the highest level of welfare is achieved by a country when it forms a TA of four members. This continues to be true for the model of heterogeneous countries for $d \in [0, d'')$, i.e. when transport costs are small. When transport costs are larger, but still in the range defined by Assumption 1, that is $d \in (d'', (e-c)/22)$, a country does better by forming a regional TA only with members from its own region. In either case, to maximize welfare, the country must form a TA with all of its regional partners.

Propositions 5 and 6 will be important in the characterization of the equilibrium TA structure with transport costs. We will see that when transport costs are low, each country would like to form a TA with a total of four countries. But because TA formation is simultaneous, each country has no way to coordinate with other countries on the TA formation process. However, when transport costs are sufficiently large, Propositions 5 and 6 taken together suggest that only regional TAs emerge in equilibrium. By Proposition 6, members of a regional TA would like a fourth country from the other region to join their TA. But by Proposition 5, no country wants to join a four-country TA that does not include all members of its own region. As a result, countries form into regional TAs. This outcome will be formalized as an equilibrium. First, I specify the game of TA formation.

4 TAs and Regions

This section sets out the game of TA formation, where countries are organized into regions. Gains from production and trade through the formation of TAs give a concrete source of payoffs to a game of non-cooperative coalition formation. The TA formation game adapts the $\delta-$ coalition formation game of Hart and Kurz (1983). The extensive form of the TA formation game is as follows. Hart and Kurz (1983) present a static 'simultaneous move exclusive membership game' in which players simultaneously announce lists of players with which they would like to form a coalition. The intersecting sets formed by the lists then form coalitions. Under their $\gamma -$ game, if a player leaves a coalition then the coalition breaks down and all players in that coalition become singletons. In their $\delta -$ game, if a player leaves a coalition then the remaining players go on to form the coalition. Hart and Kurz study the *Nash strong* equilibria of this game. A coalition structure is a Nash strong equilibrium if and only if no coalition from the set of players can improve upon the welfare of at least one member of the new coalition, with no member of the new coalition being made worse off.

I use this basic structure of the $\delta -$ game but adapt it in the following ways. First, the game is dynamic in the present setting. This makes it possible to ask whether or not the process of TA formation is consistent with a movement towards world free trade. Second, an alternative concept of equilibrium is borrowed from Arnold and Wooders (2005); that of a *Nash club* equilibrium in which, within a period, deviations may only take place within an existing TA. This arguably better reflects the process of TA formation in a world where countries are better informed about the efforts of their own TA to form than those of other countries trying to form other TAs.

4.1 The TA Formation Game

The game lasts three periods; $t = 0, 1, 2$. The process is initialized at $t = 0$ with a TA structure in which there are *no TAs*; initially, the TA structure, B, is the set of singletons. Within each period $t \geq 1$, the sequence of events is as follows. At the start of the period, each country observes the TA structure of the previous period. Then, each country i chooses a *strategy* s_i, where each s_i contains a list of countries in N with which country i would like to form a TA in the current period (this list includes country i itself). The *strategy space* S_i for country i is the set of all subsets of N i.e. the set of all possible TAs that could include country i. Strategies are chosen simultaneously. During the TA formation process, a country only observes whether or not it ends up in a TA and, if so, it sees which other countries are its TA partners. A country does not observe the strategies of other countries. We will say that, during the TA formation process, if a country does not observe another country as its TA partner, it maintains the assumption that the trade policy of that other country is described by the TA structure B of the previous period.

A *bilateral trade accord* (i, j) is formed if and only if $i \in s_j$ and $j \in s_i$. A subset of countries B_k is a *TA* if and only if all pairs of countries in B_k have a bilateral trade accord. The purpose of including i in s_i is that then we can view B_k as the intersecting set of all the elements of strategies s_i for all $i \in N$. This assumption ensures that a TA forms if and only if there is unanimous support for its membership. If a country finds itself in the position of being in two or more otherwise exclusive and otherwise unanimous TAs, it chooses the TA that maximizes its payoff under the assumption that the memberships of the TA it joins and the TA that it leaves remains otherwise constant.[18] When a country chooses one TA over another one, it assumes that the other goes ahead without it. If a new TA forms by the merger of more than one existing TA, then all members of all existing TAs must agree to the new one.

Under the assumption that countries observe the TA structure given by B in the previous period and take this as given, it is not possible to break up an existing TA in the process of forming a new one. Therefore, the assumption introduces a degree of inertia into the formal characterization of existing TAs. Countries are unable to force out existing TA partners once a TA has formed. In one sense this is theoretically restrictive, but it reflects actual practical restrictions on the cessation arrangements of existing trade agreements. For example, with regard to the EU, any member of the council of ministers has the power to veto membership of a country that would like to join, but there is no way to force out a country that is already a member. The present formalization reflects perfectly this type of arrangement.

Each strategy vector gives rise to a unique coalition structure $B(s)$:

$$B(s) = \{ (i,j) | \, i \in s_j, j \in s_i \} .$$

The payoff to country i associated with s is simply $w_i = w(t_{ir}(B(s)), t_{inr}(B(s)); d)$; the payoff for the TA structure *induced by* s. For compactness, we may write $w_i = w(s)$.

The notion of equilibrium is adapted from Arnold and Wooders (2005). For any given TA structure $B = (B_1, B_2, ..., B_k, ..., B_m)$, a strategy vector $s^* \in S$ is a *Nash club equilibrium* of the TA formation game if for any given $B_i \in B$ there is no $Z \subseteq B_i$ and $s \in S$ such that

1. $s_i = s_i^*$ for all $i \notin Z$.

[18] Pushing this one step further, any two countries caught between two CUs will assume that each behaves in the same way as the other in the CU that they choose. This assumption is the same as that of Hart and Kurz, that if any player is caught between two coalitions then it chooses the biggest one under the assumption that all other players caught in the same situation do the same. In a symmetrical world this assumption is innocuous. In principle this assumption could lead to mistakes in an asymmetrical world but this potential problem will not be an issue for any of the situations that I study.

2. $w(s_i) \geq w(s_i^*)$ for all $i \in Z$ and $w(s_i) > w(s_i^*)$ for some $i \in Z$.

It is worth emphasizing that by the definition of equilibrium, deviations are only allowed by a group of countries within a TA, $Z \subseteq B_k$. This is in keeping with equilibrium as defined by Arnold and Wooders (2005), who similarly restrict communication and hence allowable deviations.

By contrast, Hart and Kurz allow deviations to be undertaken by *any* coalition of countries $Z \subseteq N$. If I were to allow deviations by any coalition of countries $Z \subseteq N$ then equilibrium may fail to exist, for reasons that shall become clear. However, Arnold and Wooders (2005) argue that in many settings the flow of information is restricted to the set of individuals that form a club. This seems to be a reasonable assumption in the context of TA formation, at least in the short run. This weakens the notion of equilibrium relative to Hart and Kurz, admitting a relatively large number of equilibria. In particular, it does not exclude from the equilibrium set candidates that arise as a result of coordination failure - in the present context, where countries could all benefit from moving to free trade but fail to do so. As we shall see, two interesting features of equilibrium arise as a result. First, while a large number of equilibria are possible in the absence of distance between countries, when distance is introduced then the greater benefit that results from a regional agreement provides a coordination device which may give rise to a unique equilibrium path. Second, this equilibrium path converges to world free trade, which is efficient, but only gradually after a period in which two regional blocks form.

One more remark is in order. The assumption of naive best responses gives a convenient and tractable way to define payoffs under the TA formation game. Naive best responses are also believed to capture the process by which agents learn about their environment when they do not have full information about, or cannot communicate perfectly with, the actions of all of the other players (see Bala and Goyal 2000). Such an assumption seems reasonable in the present international policy making environment where policy makers are not able to perfectly observe each others' actions. An alternative approach would be to assume that countries are far sighted. That is, we could assume that countries can communicate perfectly and can rationally anticipate the incentives of other groups to form trade agreements, and can take this into account when proposing their own trade agreement membership.[19] Page,

[19]Perfect communication appears to be a necessary counterpart to farsightedness in the present model because there are so many agreements that a country could anticipate forming. In the absence of commu-

Wooders and Kamat (2004) consider a general network formation game in which players are farsighted.

5 No Transport Costs; The Problem of Coordination Failure

5.1 Various Equilibria with Coordination Failure

Fix $d = 0$. By Proposition 6 we know that a TA of four countries maximizes the welfare of its members (if the other two countries are singletons). The problem of coordination failure arises because even if each country writes down a strategy s_i with four elements, aiming to form a TA with four countries, in the absence of communication there are many possible TA structures that may arise in equilibrium as a result. An equilibrium may arise in which there is an agreement with four countries, which is the desired outcome of each of the members. But of course the two excluded countries do not achieve their desired objective from the strategy that they play. Each is in an agreement of two countries (i.e. with each other) and not four as they intended. Moreover, this is not the only trade agreement structure that can be sustained in equilibrium. We will first consider an equilibrium in which there is a single four-country agreement, but then consider one of many possible alternative trade agreement structures that may arise.

An example of a strategy vector that gives rise to an equilibrium in which there is a four-country agreement is as follows:

$$\begin{aligned} s_1 &= \{1,2,3,4\} \\ s_2 &= \{1,2,3,4\} \\ s_3 &= \{1,2,3,4\} \\ s_4 &= \{1,2,3,4\} \\ s_5 &= \{1,2,3,5\} \\ s_6 &= \{1,2,3,6\}. \end{aligned}$$

Notice that the strategies $s_1 ... s_4$ form an intersecting set of elements $\{1, 2, 3, 4\}$ while the element 5 is not contained in any strategy other than s_5 and 6 is not contained in any strategy

nication, it is unclear how a country would work out which agreement it should form, given that all other agents are far sighted.

other than s_6. Thus, the resulting trade agreement structure is $\{\{1,2,3,4\},\{5\},\{6\}\}$. It is easy to check that no country can gain by deviating from this agreement structure and so therefore this must be an equilibrium. Consider the allowable deviations. If a member of the four-country agreement were to veto membership of another single member then the agreement structure would become one of a three-country agreement and three singletons, for example $\{\{1,2,3\},\{4\},\{5\},\{6\}\}$. Then, by Proposition 6, the payoff to the country that undertook the veto would fall, as would the payoff of the ejected member. The welfare of the singletons $\{5\}$ and $\{6\}$ actually increases. If more than one country's membership is vetoed, it is easy to check that the payoff of remaining members falls even further. No member of the four-country agreement has an incentive to deviate. No deviations are available to the singletons. Thus we have a Nash club equilibrium.

We have already discussed above the reasons why agreement member welfare changes when one or more countries are ejected from the agreement. Let us briefly review why non-member welfare changes. We just noted that, from an initial trade agreement structure of $\{\{1,2,3,4\},\{5\},\{6\}\}$, if Country 4 is ejected, leaving a trade agreement structure of $\{\{1,2,3\},\{4\},\{5\},\{6\}\}$, then the welfare of 5 and 6 increases. Why does this happen? Tariffs set by 5 and 6 do not change because these depend only on their own trade agreement structure, which has not changed. When 4 is ejected, Countries 1, 2 and 3 restore tariffs against it, and as a result demand less from 4, shifting some of their demand at the margin towards 5 and 6. This increases profits in 5 and 6. In addition, 4 restores tariffs against Countries 1, 2 and 3, shifting its demand towards 5 and 6. Both of these effects combine to shift profits towards 5 and 6, thus increasing welfare.

Notice that the partition of countries into regions has no relevance to this equilibrium. As specified, the equilibrium contains three countries from R_1 and one country from R_2. But under and equivalent characterization of equilibrium we could permute the countries in such a way that two countries were in R_1 (say 1 and 2), and two countries were in R_2 (say 3 and 4). This is due to the fact that all countries are symmetrical, which follows from the present assumption that $d = 0$. We shall see that the partition of countries into regions is relevant for equilibrium when $d > 0$.

Now let us consider another possible equilibrium, in which there are three agreements each with two members. This equilibrium arises if each country proposes to form a TA with the three countries ‘next to it’:

$$\begin{aligned} s_1 &= \{1,2,3,4\} \\ s_2 &= \{2,3,4,5\} \\ s_3 &= \{3,4,5,6\} \\ s_4 &= \{4,5,6,1\} \\ s_5 &= \{5,6,1,2\} \\ s_6 &= \{6,1,2,3\}. \end{aligned}$$

By inspection of the strategy vector, the agreements that form are $\{1,4\}$, $\{2,5\}$ and $\{3,6\}$. Again, it is straight forward to check that this is an equilibrium strategy vector. If any member of a two-country agreement vetoes membership of the other, splitting the agreement into two singleton agreements, then its payoff falls by Proposition 3. This is the only feasible deviation. Consequently, the strategy vector shown above must be an equilibrium.

Again, notice that the partition of countries into regions has no relevance to this equilibrium. As specified, each two-country agreement comprises one country from each region. But, again, a permutation of countries yields an equilibrium in which one agreement has two countries from one region, one agreement has two countries from the other region and one agreement has one country from each region.

6 Transport Costs and Coordination

6.1 Coordination Problem Resolved with Transport Costs

The problem of coordination failure identified in the previous section is resolved in the presence of transport cost $d \in (0, (e-c)/22)$. It follows from Proposition 6 that in period $t = 1$ each country has an incentive to form a trade agreement with all other countries in the same region. With transport costs $d > d''$, each country also wants to include one country from the other region in its TA.

Proposition 7. *Assume $d \in (0, (e-c)/22)$. At $t = 1$ there is a unique equilibrium with two regional TAs; $B_1 = R_1$, $B_2 = R_2$. The payoff to each country is the same and is lower than free trade.*

There are two cases to consider, although the outcome is the same in both; one where $d \in (0, d'')$ and one is where $d \in (d'', (e-c)/22)$. For the case where $d \in (d'', (e-c)/22)$ the outcome is obvious; by Proposition 6, each country does better in a regional TA than in any other TA, given that all other countries are singletons. So it is immediate that the intersecting sets formed by countries' strategies is two regional TAs; $B_1 = R_1$ and $B_2 = R_2$.

The case where $d \in (0, d'')$ is slightly more subtle, but the outcome is the same. In that case, each country's welfare is maximized by a 4-member TA with three members from its own region and one member from the other region. But even if all countries write down a strategy containing four countries, three from its own region and one from the other region, the intersecting sets of countries formed by these strategies give rise to two regional TAs; $B_1 = R_1$ and $B_2 = R_2$.

Consider, for example, the following strategy vector:

$$\begin{aligned} s_1 &= \{1,2,3,4\} \\ s_2 &= \{1,2,3,4\} \\ s_3 &= \{1,2,3,4\} \\ s_4 &= \{1,4,5,6\} \\ s_5 &= \{1,4,5,6\} \\ s_6 &= \{1,4,5,6\}. \end{aligned}$$

The strategies $s_1 ... s_3$ form an intersecting set of elements $\{1,2,3\}$ and the strategies $s_4 ... s_6$ form an intersecting set of elements $\{4,5,6\}$. Thus, the resulting trade agreement structure is $\{\{1,2,3\},\{4,5,6\}\}$. It is easy to check that no country can gain by deviating from this agreement structure and so therefore this must be an equilibrium. In the proposed equilibrium the trade agreement structure is symmetrical, so each country receives the same payoff. By Proposition 4, the payoff that each country receives must be lower than under free trade.

Consider the allowable deviations. If a member of one of the regional agreements were to veto membership of another single member then the agreement structure would become one of a two-country agreement, a singleton and a three-country agreement; for example $\{\{1,2,3\},\{4,5\},\{6\}\}$. Then the payoff to the country that undertook the veto, in this example country 4, would fall. The welfare of countries in the regional trade agreement that remains $\{1,2,3\}$ actually increases. As before, if more than one country's membership is vetoed, it is easy to check that the payoff of the remaining member falls even further. Thus, no member of a regional agreement has an incentive to deviate. No deviations are available to the singletons. Thus we have a Nash club equilibrium. Providing play proceeds in the manner described, this is the only possible equilibrium that can arise for transport costs in the interval $d \in (0, (e-c)/22))$.

Clearly, the assumption that agents hold constant the strategies of other countries when forming a trade agreement is crucial for this outcome. If countries were far-sighted then each would obviously anticipate that the countries of the other region would form a trade agreement as well. Then a move to free trade would be more beneficial. Other outcomes might be supported too under farsightedness, depending on whether any other information becomes available through farsightedness. But we can also see how the present assumption of naivete captures aspects of uncertainty that are almost certainly present in the actual process of trade agreement formation across regions.

7 Do Regional TAs Facilitate Free Trade?

We have seen how two regional trade agreements emerge at stage $t = 1$. We now proceed to stage $t = 2$ and ask whether free trade can emerge at this point. We find that it does. The thinking is as follows. Each country observes the regional trading arrangements described by B from period $t = 1$. From $t = 1$, there are two regional trade agreements; $B_1 = R_1$, $B_2 = R_2$. At $t = 2$, countries are able to secure this same payoff as at $t = 1$ by maintaining the existing structure. However, each is able to obtain a higher payoff by moving to free trade. Thus every country is potentially able to gain by moving to free trade.

Proposition 8. *Assume $d \in \{0, (e-c)/22\}$. There is a unique equilibrium path. At $t = 1$ there are two regional TAs; $B_1 = R_1$, $B_2 = R_2$. At $t = 2$ there is world free trade.*

The reference to $t = 1$ replicates Proposition 1 and is included for completeness. To check equilibrium at $t = 2$, we need to check incentives to deviate. Also, following the extensive form specification of the game we need to check that each country i would write down a strategy, s_i, that would bring about free trade. These tasks will be taken in turn.

At $t = 2$ there is no incentive for any country $i \in N$ to deviate from free trade. This contrasts with the situation at $t = 1$. At $t = 1$, country i could obtain its highest level of welfare by forming a regional TA while all countries in the other region remained as singletons. And, by Proposition 6, the payoff is higher than free trade. How can free trade be an equilibrium at $t = 2$ but not at $t = 1$? At $t = 1$ a country has an incentive to veto the membership of the three countries in the other region. The payoff to such a deviation rests on the assumption that all the ejected members return to the trade agreement structure given by the network B at $t = 0$. That is, all countries were assumed to return to singleton status. At $t = 2$ the outcome is different. All ejected countries are assumed to return to the trade agreement structure given by B at $t = 1$; a regional trade agreement. Thus, deviation by ejection of countries from the other region returns the trade agreement structure to one of two regional agreements. By Proposition 7, the payoff to such a deviation is not profitable as it is lower than free trade. Also notice that, by assumption, it is not possible to break apart an existing trade agreement by ejecting a subset of countries from the other region. If a deviation from free trade were allowed in which only one of the countries were ejected, for example bringing about a trade agreement structure of $\{\{1,2,3,4,5\},\{6\}\}$ then countries 1, 2 and 3 may be able to gain over free trade.

Let us now look at the incentive for country i to play the strategy $s_i = \{1,2,3,4,5,6\}$. Obviously the payoff to free trade is higher than the status quo given by the strategy $s_i = \{1,2,3\}$. As for the deviation considered above, it may appear that a higher payoff could be obtained for countries $1-5$ by writing down a strategy $s_i = \{1,2,3,4,5\}$ for example. But again, by assumption, when a new TA is formed from two existing trade agreements it must contain all countries of the two original agreements. Thus a strategy $s_i = \{1,2,3,4,5\}$ is not allowable at $t = 2$, given the equilibrium $\{\{1,2,3\},\{4,5,6\}\}$ at $t = 1$. Indeed, the only strategy that includes all the countries from the two agreements of period $t = 1$ (and that yields an increase in welfare) is the strategy yielding free trade. Thus, we have shown that free trade is a Nash club equilibrium at $t = 2$. Regional trade agreements do ultimately facilitate free trade.

8 Conclusions

The purpose of this paper has been to show, using an example, that problems of coordination failure in the formation of regional trade agreements may be resolved when countries are organized into regions. Costs of shipping goods between regions must be significant, but no so high as to eliminate trade between regions. With no transport costs, there is a problem of multiple equilibria due to coordination failure familiar from the theory of networks. Positive transport costs (at a sufficiently high level) are enough to bring about a unique equilibrium in the first period of the trade agreement network formation game. Starting from a situation where there are no trade agreements, through the coalition formation process, in the first period two regional trade agreements form simultaneously. In the second period the two regional trade agreements merge to bring about free trade. The attainment of free trade only after a period of regionalism rests on a restriction in the flow of information through the TA formation process. Members can only communicate about their agreement, once they have simultaneously and independently chosen their trade agreement partners. Best responses are made naively, by gleaning information from the TA of the previous period about other countries' policies.

The underlying economic structure of the model is one of Cournot competition. I believe that the features of the model exhibited in the examples would carry over to other forms of production. It is widely appreciated that Bertrand competition behaves like Cournot competition when firms must pre-commit to quantities. An endowment economy, or a more elaborate modelling of perfect competition should also exhibit the same features, as suggested by Bond (2001). The key motivating feature of the model is that import substitution elasticities are declining in distance in the model, and this motivates higher rents in trade and hence higher tariffs between close neighbors in the absence of an agreement. This feature of the model should be robust to alternative assumptions about production.

One question that should be addressed in future research on this topic is whether the predictions of the example are robust to non-linear demands. This type of question would almost certainly need to be addressed using simulations, since the feature of the model that optimal tariffs have analytical solutions depends on the linearities in the model. Simulation analysis could also tackle the broarder question of whether the model predictions are robust to more elaborate and realistic country and regional structures. How big must the asymmetries across countries and regions get before problems of multiple equilibria re-emerge? When will asymmetries preclude the eventual move to free trade?

Another interesting line of research is to investigate how variation in the assumptions over the flow of information and expectations between countries through the agreement formation process changes the outcome. It appears that perfect information and perfect foresight facilitate an immediate move to free trade. But it would be interesting to ask in which ways weakening information flows in various ways would vary the outcome away from free trade, and under what alternative assumptions about information flows and expectations regional trade block formation would be the result.

A Appendix

A.1 Demand functions by region and agreement membership

Based on (6), the basic expressions for equilibrium output produced by country j for country i can be written as follows. The output functions (8)-(11) are obtained by substituting optimal tariffs $t_{ir}^*(b_{ir}, b_{inr})$ and $t_{inr}^*(b_{ir}, b_{inr})$ into the following functions:

$$x_{irm} = \frac{(e-c) + 3d + (3-b_{ir})\,t_{ir} + (3-b_{inr})\,t_{inr}}{7}.$$

$$x_{irnm} = \frac{(e-c) + 3d - (4+b_{ir})\,t_{ir} + (3-b_{inr})\,t_{inr}}{7}.$$

$$x_{inrm} = \frac{(e-c) - 4d + (3-b_{ir})\,t_{ir} + (3-b_{inr})\,t_{inr}}{7}.$$

$$x_{inrnm} = \frac{(e-c) - 4d + (3-b_{ir})\,t_{ir} - (4+b_{inr})\,t_{inr}}{7}.$$

A.2 Proofs

Proof of Proposition 1.

Maximize the function for w_i, (7), with respect to t_{ir} and t_{inr} using $C_i = X_i^2/2$ and $T_i = \sum_{k\in N} t_{ik}x_{ik}$, and (6). The proof that (7) is concave in t_{ir} and t_{inr} respectively follows Yi (1996).

From the first order condition for the maximization of w_i with respect to t_{ir}, using (8)-(11) and rearranging, it is possible to obtain a reduced form expression for t_{ir}^* in terms of all other variables:

$$t_{ir}^* = \frac{(1+2(b_{ir}+b_{inr}))(e-c) + 2(12+3b_{ir}-4b_{inr})\,d - (3-b_{inr})(15+2(b_{ir}+b_{inr})\,t_{inr})}{53-2b_{inr}(3-b_{ir}) + b_{ir}(9+2b_{ir})}$$

In the same way, it is possible to obtain a reduced form expression for t_{inr}^* in terms of all other variables:

$$t_{inr}^* = \frac{(1+2(b_{ir}+b_{inr}))(e-c) - (25-6b_{ir}+8b_{inr})\,d + (3-b_{ir})(15+2(b_{ir}+b_{inr})\,t_{ir})}{53-6b_{ir}+b_{inr}(9+2(b_{ir}+b_{inr}))}.$$

Solving simultaneously for t_{ir}^* and t_{inr}^* obtains the result. □

Proof of Lemma 2. That fact that $x_{irm}(b_{ir}, b_{inr}) > x_{inrm}(b_{ir}, b_{inr}) > x_{irnm}(b_{ir}, b_{inr}) > x_{inrnm}(b_{ir}, b_{inr})$ is established using (8)-(11), by inspection or otherwise. It remains to show that if $d \in (0, (e-c)/22)$ then $x_{inrnm}(b_{ir}, b_{inr}) > 0$. Since $x_{inrnm}(b_{ir}, b_{inr})$, as given by (11), is decreasing d, we can solve for the largest value of d at which $x_{inrnm}(b_{ir}, b_{inr}) = 0$. The solution is

$$d = \frac{2(e-c)}{5 + 3b_{ir} + 2b_{ir}b_{inr} + 2b_{ir}^2}.$$

Finding the feasible pair $\{b_{ir}, b_{inr}\}$, $b_{ir} \in \{1,2,3\}$, $b_{inr} \in \{0,1,2\}$ at which d is minimized will find the value of d below which $x_{inrnm}(b_{ir}, b_{inr}) > 0$. (Recall that there are no non-regional non-members for $b_{ir} = 3$, $b_{inr} = 3$.) The solution is globally decreasing in b_{ir} and b_{inr}, so use $b_{ir} = 3$, $b_{inr} = 2$ in the solution to yield $d = (e-c)/22$. The result follows. □

Proof of Proposition 2. It remains only on establish that $dx_{irnm}/db_{ir} < 0$, $dx_{inrnm}/db_{ir} < 0$, $dx_{irnm}/db_{inr} < 0$, and $dx_{inrnm}/db_{inr} < 0$. Each case will be taken in turn. Differentiating $x_{irnm}(b_{ir}, b_{inr})$ with respect to b_{ir}, we obtain

$$\begin{aligned} \frac{dx_{irnm}(b_{ir}, b_{inr})}{db_{ir}} &= -\frac{(3 + 4b_{ir} + 2b_{inr})\, d}{2\Delta(b_{ir}, b_{inr})} \\ &\quad -\frac{(6 + 8(b_{ir} + b_{inr}))(2(e-c) + (3 + b_{inr}(3 + 2(b_{ir} + b_{inr})))\, d)}{(2\Delta(b_{ir}, b_{inr}))^2} \end{aligned}$$

$$= -\frac{2(3 + 4(b_{ir} + b_{inr}))(e-c) + (9 + 12b_{ir} + b_{inr}(1 + 2(b_{ir} + b_{inr}))(5 + 2(b_{ir} + b_{inr})))\, d}{2(\Delta(b_{ir}, b_{inr}))^2}$$

So $dx_{irnm}(b_{ir}, b_{inr})/db_{ir} < 0$ for all $d \geq 0$.

Differentiating $x_{irnm}(b_{ir}, b_{inr})$ with respect to b_{inr}, we obtain

$$\begin{aligned} \frac{dx_{irnm}(b_{ir}, b_{inr})}{db_{ir}} &= -\frac{2b_{ir}d}{2\Delta(b_{ir}, b_{inr})} \\ &\quad -\frac{(6 + 8(b_{ir} + b_{inr}))(2(e-c) + (3 + b_{inr}(3 + 2(b_{ir} + b_{inr})))\, d)}{(2\Delta(b_{ir}, b_{inr}))^2} \end{aligned}$$

$$= \frac{-2(3 + 4(b_{ir} + b_{inr}))(e-c) + (15 + 13b_{ir} + 4(5b_{inr} + b_{ir}b_{inr}(3 + b_{inr}) + (3 + 2b_{inr})b_{ir}^2 + b_{ir}^3))\, d}{2(\Delta(b_{ir}, b_{inr}))^2}$$

Again, the second term in the numerator is positive and increasing in b_{ir}, b_{inr} and d while the first term is negative. (It is easily checked that overall the numerator is negative for $b_{ir} = 3$, $b_{inr} = 2$ and $d = (e-c)/22$. So $dx_{irnm}(b_{ir}, b_{inr})/db_{inr} < 0$ for all feasible $\{b_{ir}, b_{inr}\}$ pairs and $d \in (e-c)/22 \geq 0$.

Differentiating $x_{inrnm}(b_{ir}, b_{inr})$ with respect to b_{ir}, we obtain

$$\frac{dx_{inrnm}(b_{ir}, b_{inr})}{db_{ir}} = \frac{2b_{inr}d}{2\Delta(b_{ir}, b_{inr})} - \frac{(6+8(b_{ir}+b_{inr}))(2(e-c)+(3+b_{inr}(3+2(b_{ir}+b_{inr})))d)}{(2\Delta(b_{ir}, b_{inr}))^2}$$

$$= -\frac{2(3+4(b_{ir}+b_{inr}))(e-c)+(9+12b_{ir}+b_{inr}(1+2(b_{ir}+b_{inr}))(5+2(b_{ir}+b_{inr})))d}{2(\Delta(b_{ir}, b_{inr}))^2}$$

After simplification, we see that $dx_{inrnm}(b_{ir}, b_{inr})/db_{ir} = dx_{irnm}(b_{ir}, b_{inr})/db_{ir} < 0$ for all $d \geq 0$.

Differentiating $x_{inrnm}(b_{ir}, b_{inr})$ with respect to b_{inr}, we obtain

$$\frac{dx_{inrnm}(b_{ir}, b_{inr})}{db_{ir}} = \frac{(3+2b_{ir}+4b_{inr})d}{2\Delta(b_{ir}, b_{inr})} - \frac{(6+8(b_{ir}+b_{inr}))(2(e-c)+(3+b_{inr}(3+2(b_{ir}+b_{inr})))d)}{(2\Delta(b_{ir}, b_{inr}))^2}$$

$$= \frac{-2(3+4(b_{ir}+b_{inr}))(e-c)+(15+13b_{ir}+4(5b_{inr}+b_{ir}b_{inr}(3+b_{inr})+(3+2b_{inr})b_{ir}^2+b_{ir}^3))d}{2(\Delta(b_{ir}, b_{inr}))^2}$$

After simplification, we see that $dx_{inrnm}(b_{ir}, b_{inr})/db_{ir} = dx_{irnm}(b_{ir}, b_{inr})/db_{ir}$ So it must be the case that $dx_{inrnm}(b_{ir}, b_{inr})/db_{inr} < 0$ for all feasible $\{b_{ir}, b_{inr}\}$ pairs and $d \in (e-c)/22 \geq 0$. □

Proposition 3. *The expansion or formation of a TA increases the aggregate welfare of member countries.*

Proof of Proposition 3. The proof follows Yi (1996). Assume that there exists a TA structure $B = (B_1, B_2, ..., B_m)$ and that two or more TAs B_1, B_2 ... B_r merge to create an enlarged TA. We will show that the total welfare of the members of the enlarged TA increases. We will do this by showing that the tariff changes required to implement TA enlargement undertaken by any one given member of the enlarged TA must increase the aggregate welfare of the enlarged TA members. Thinking of TA enlargement as a sequence of such tariff changes by each and every member then gives the result.

Claim. Initially, before the merger, country i has free trade with $b_{ir} - 1$ countries in its own region and b_{inr} countries in the other region. Country i levies a tariff $t_{ir}(b_{ir}, b_{inr})$ on each of the $3 - b_{ir}$ non-members in its own region and a tariff $t_{inr}(b_{ir}, b_{inr})$ on each of the $3 - b_{inr}$ countries in the other region. As a result of the merger, in the new enlarged

TA, country i shares a TA with $b'_{ir}-1$ countries in its own region and b'_{inr} countries in the other region. Let $s_{ir}=b'_{ir}-b_{ir}\geq 0$ and $s_{inr}=b'_{inr}-b_{inr}\geq 0$. Country i abolishes tariffs on s_{ir} countries in its own region and s_{inr} countries in the other region, and changes tariffs to $t'_{ir}\left(b'_{ir},b'_{inr}\right)$ on each of the $3-b'_{ir}$ non-members in its own region and changes tariffs to $t'_{inr}\left(b'_{ir},b'_{inr}\right)$ on each of the $3-b'_{inr}$ non-members in the other region. Then the aggregate welfare of the $b_{ir}+s_{ir}+b_{inr}+s_{inr}$ countries in the enlarged TA (which consists of country i, $b_{ir}+b_{inr}$ countries who paid no tariffs initially and $s_{ir}+s_{inr}$ countries whose tariffs were abolished) improves.

Proof. Without loss of generality, consider the TA B_1, of which Country 1 is assumed to be a member. The cardinality of B_1 is $b_{1r}+b_{1nr}$, with b_{1r} members from R_1 and b_{1nr} members from R_2. Then let membership expand to create an enlarged TA, B'_1 consisting of b'_{1r} members in R_1 and b'_{1nr} members in R_2 (where all original members are also members of the enlarged TA). The comparative statics exercise that we will now carry out is as follows. We will calculate the effect on the aggregate welfare of all countries in B'_1 that results when Country 1 abolishes tariffs on s_{1r} countries in R_1 and s_{1nr} countries in R_2, and changes tariffs on $(3-b_{ir}-s_{ir})$ non-members in R_1 from $t_{1r}\left(b_{1r},b_{1nr}\right)$ to $t_{1r}\left(b'_{1r},b'_{1nr}\right)$ and on $(3-b_{inr}-s_{inr})$ non-members in R_2 from $t_{1nr}\left(b_{1r},b_{1nr}\right)$ to $t_{1nr}\left(b'_{1r},b'_{1nr}\right)$.

Define

$$\begin{aligned}\Delta t_{ir} &= t_{ir}\left(b_{ir},b_{inr}\right)-t_{ir}\left(b'_{ir},b'_{inr}\right);\\ \Delta t_{inr} &= t_{inr}\left(b_{ir},b_{inr}\right)-t_{inr}\left(b'_{ir},b'_{inr}\right).\end{aligned}$$

First consider infinitesimal changes in tariffs

$$d\mathbf{t}\equiv\left(0,...,0,dt,...,dt,dt_r,...,dt_r,0,...,0,\phi dt,...,\phi dt,dt_{nr},...,dt_{nr}\right)$$

from a tariff vector

$$\mathbf{t}\equiv\left(0,...,0,t,...,t,t_r,...,t_r,0,...,0,\phi t,...,\phi t,t_{nr},...,t_{nr}\right),$$

where: dt appears from the $(b_{ir}+1)$th element to the $(b_{ir}+s_{ir})$th element and from the $(b_{inr}+4)$th element to the $(b_{inr}+4+s_{inr})$th element, unless $b_{inr}=s_{inr}=0$ in which case dt_{nr} appears from the 4th to the last elements; dt_r appears from the $(b_{ir}+s_{ir}+1)$th element to the 3rd element; dt_{nr} appears from the $(b_{inr}+s_{inr}+4)$th element to the last element. The tariff t is imposed on new TA members in the same region and is reduced to zero through the TA formation process. The term ϕt is the tariff imposed on new TA members from the other region, where ϕ measures the ratio of non-regional to regional tariffs (to be defined below). Also,

$$\begin{aligned} dt_r &\equiv \frac{\Delta t_{ir}}{t_{ir}\left(b_{ir}, b_{inr}\right)} dt; \\ dt_{nr} &\equiv \frac{\Delta t_{ir}}{t_{inr}\left(b_{ir}, b_{inr}\right)} \phi dt. \end{aligned}$$

Start from

$$\mathbf{t}\left(b'_{ir}, b'_{inr}\right) \equiv \left(0, ..., 0, t_{ir}\left(b'_{ir}, b'_{inr}\right), ..., t_{ir}\left(b'_{ir}, b'_{inr}\right), 0, ..., 0, t_{inr}\left(b'_{ir}, b'_{inr}\right), ..., t_{inr}\left(b'_{ir}, b'_{inr}\right)\right)$$

where 0 appears from the first to the $(b_{ir} + s_{ir})$th element and from the fourth to the $(b_{inr} + s_{inr} + 4)$th element (unless $b_{inr} = s_{inr} = 0$). We can move to

$$\mathbf{t}\left(b_{ir}, b_{inr}\right) \equiv \left(0, ..., 0, t_{ir}\left(b_{ir}, b_{inr}\right), ..., t_{ir}\left(b_{ir}, b_{inr}\right), 0, ..., 0, t_{inr}\left(b_{ir}, b_{inr}\right), ..., t_{inr}\left(b_{ir}, b_{inr}\right)\right)$$

where 0 appears from the first to the (b_{ir})th element and from the fourth to the $(b_{inr} + 4)$th element by integrating the infinitesimal changes $d\mathbf{t}$ from 0 to $\mathbf{t}\left(b_{ir}, b_{inr}\right)$. Below, we will show that $d\left(\sum_{k\in B'_1} w_k\right)/d\mathbf{t} < 0$ for all $\mathbf{t}$ along such a path of integration. The claim then follows.

Since changes in Country 1's tariffs do not affect sales in other countries,

$$d\left(\sum_{k\in B'_1} w_k\right)/d\mathbf{t} = d\left(\hat{w}_1 + \sum_{k\in B'_1\setminus\{1\}} \pi_{ik}\right)/d\mathbf{t}$$

where $\hat{w}_1$ is Country 1's welfare net of its exports. Since $\hat{w}_1 + \sum_{k\in N} \pi_{1k} = v\left(X_1\right) - cX_1$, $\hat{w}_1 + \sum_{k\in B'_1\setminus\{1\}} \pi_{1k} = v\left(X_1\right) - cX_1 - \sum_{k\in N\setminus B_1} \pi_{1k}$.

The proportional relationship between $t_{ir}\left(b_{ir}, b_{inr}\right)$ and $t_{inr}\left(b_{ir}, b_{inr}\right)$ is given by

$$\begin{aligned} \phi &= \frac{t_{inr}\left(b_{ir}, b_{inr}\right)}{t_{ir}\left(b_{ir}, b_{inr}\right)} \\ &= 1 - \frac{2\left(4 + 5b_{inr} + 2\left(b_{inr} - 1\right)b_{ir} + 2b_{ir}^2\right)d}{\left(1 + 2\left(b_{ir} + b_{inr}\right)\right)\left(e - c\right) + \left(3 + b_{ir}\left(2\left(b_{ir} + b_{inr}\right) - 1\right)\right)d}. \end{aligned}$$

The total tariff at the tariff vector $\mathbf{t}$ is

$$T_i = \sum_{k\in N} t_{ik} = \left(s_{ir} + \phi s_{inr}\right)t + \left(3 - b_{ir} - s_{ir}\right)t_r + \left(3 - b_{inr} - s_{inr}\right)t_{nr},$$

The change in the total tariff is calculated from $d\mathbf{t}$ as follows:

$$\begin{aligned} dT_i &= \left(s_{ir} + \phi s_{inr}\right) dt + \left(3 - b_{ir} - s_{ir}\right) dt_r + \left(3 - b_{inr} - s_{inr}\right) dt_{nr} \\ &= \frac{s_{ir} t_{ir}\left(b_{ir}, b_{inr}\right) + s_{inr} t_{inr}\left(b_{ir}, b_{inr}\right) + \left(3 - b_{ir} - s_{ir}\right) \Delta t_{ir} + \left(3 - b_{inr} - s_{inr}\right) \Delta t_{inr}}{t_{ir}\left(b_{ir}, b_{inr}\right)} dt. \end{aligned}$$

The following notation will also be helpful:

$$\Delta T_i = s_{ir} t_{ir}\left(b_{ir}, b_{inr}\right) + s_{inr} t_{inr}\left(b_{ir}, b_{inr}\right) + \left(3 - b_{ir} - s_{ir}\right) \Delta t_{ir} + \left(3 - b_{inr} - s_{inr}\right) \Delta t_{inr}$$

From (4) and the first-order-condition of (5), $p_i - c = x_{ij} + t_{ij} + d_{ij}$. Therefore, $\sum_{k \in N} \left(p_i - c\right) = X_i + T_i + D_i$, where $D_i = \sum_{k \in N} d_{ik} = 3d$.

From (6), $dx_{ij} = \frac{dT_i - 7dt_j}{7}$. Therefore we have:

$$\begin{aligned} \frac{dx_{ii}}{d\mathbf{t}} &= \frac{\Delta T_i}{7t_{ir}\left(b_{ir}, b_{inr}\right)}; \\ \frac{dx_{ib_{ir}+1}}{d\mathbf{t}} &= \frac{\Delta T_i - 7t_{ir}\left(b_{ir}, b_{inr}\right)}{7t_{ir}\left(b_{ir}, b_{inr}\right)}; \\ \frac{dx_{ib_{ir}+s_{ir}+1}}{d\mathbf{t}} &= \frac{\Delta T_i - 7\Delta t_{ir}}{7t_{ir}\left(b_{ir}, b_{inr}\right)}; \\ \frac{dx_{ib_{inr}+4}}{d\mathbf{t}} &= \frac{\Delta T_i - 7t_{inr}\left(b_{ir}, b_{inr}\right)}{7t_{ir}\left(b_{ir}, b_{inr}\right)}; \\ \frac{dx_{ib_{inr}+s_{inr}+4}}{d\mathbf{t}} &= \frac{\Delta T_i - 7\Delta t_{inr}}{7t_{ir}\left(b_{ir}, b_{inr}\right)}. \end{aligned}$$

Using these results,

$$\begin{aligned} \frac{d}{d\mathbf{t}}\left(\hat{w}_1 + \sum_{k \in B_1' \setminus \{1\}} \pi_{ik}\right) &= \frac{d}{d\mathbf{t}}\left[v\left(X_1\right) - cX_1\right] - \frac{d}{d\mathbf{t}} \sum_{k \in N \setminus B_1'} \pi_{1k} \\ &= \sum_{k \in N} \left[p_1 - c\right] \frac{dx_{1k}}{d\mathbf{t}} - \sum_{k \in N \setminus B_1'} 2x_{1k} \frac{dx_{1k}}{d\mathbf{t}} \\ &= \frac{1}{7t_{1r}\left(b_{1r}, b_{1nr}\right)} \left\{s_{1r} t_{1r}\left(b_{1r}, b_{1nr}\right) \Xi_1 + s_{1nr} t_{1nr}\left(b_{1r}, b_{1nr}\right) \Phi_1 \right. \\ &\quad \left. + \left(3 - b_{1r} - s_{1r}\right) \Delta t_{1r} \Psi_1 + \left(3 - b_{1nr} - s_{1nr}\right) \Delta t_{1nr} \Omega_1 \right\}, \end{aligned}$$

where:

$$\begin{aligned}
\Xi_i &= (X_i + T_i + D_i) - 7\left(x_{ib_{ir}+1} + t\right) \\
&\quad -2\left(3 - b_{ir} - s_{ir}\right) x_{ib_{ir}+s_{ir}+1} - 2\left(3 - b_{inr} - s_{inr}\right) x_{ib_{inr}+s_{inr}+4}; \\
\Phi_i &= (X_i + T_i + D_i) - 7\left(x_{ib_{inr}+4} + \phi t + d\right) \\
&\quad -2\left(3 - b_{ir} - s_{ir}\right) x_{ib_{ir}+s_{ir}+1} - 2\left(3 - b_{inr} - s_{inr}\right) x_{ib_{inr}+s_{inr}+4}; \\
\Psi_i &= (X_i + T_i + D_i) - 7\left(x_{ib_{ir}+s_{ir}+1} + t_r\right) + 2\left(4 + b_{ir} + s_{ir}\right) x_{ib_{ir}+s_{ir}+1}; \\
\Omega_i &= (X_i + T_i + D_i) - 7\left(x_{ib_{inr}+s_{inr}+4} + t_{nr} + d\right) + 2\left(4 + b_{inr} + s_{inr}\right) x_{ib_{inr}+s_{inr}+4}.
\end{aligned}$$

The proof that $\frac{d}{d\mathbf{t}}\left(\hat{w}_1 + \sum_{k\in B_1'\backslash\{1\}} \pi_{ik}\right) < 0$ proceeds in two steps. First we show that, at $\mathbf{t}\left(b_{ir}', b_{inr}'\right)$, it is the case that $\frac{d}{d\mathbf{t}}\left(\hat{w}_1 + \sum_{k\in B_1'\backslash\{1\}} \pi_{ik}\right) < 0$. Second, we show that $\frac{d^2}{d\mathbf{t}^2}\left(\hat{w}_1 + \sum_{k\in B_1'\backslash\{1\}} \pi_{ik}\right) < 0$.

Step 1. At $\mathbf{t}\left(b_{ir}', b_{inr}'\right)$, the optimal tariffs $t_{ir}\left(b_{ir}', b_{inr}'\right)$ and $t_{inr}\left(b_{ir}', b_{inr}'\right)$ are chosen to satisfy $\Psi_i = 0$ and $\Omega_i = 0$ respectively. (Note that Ψ_i and Ω_i are the derivatives of $\hat{w}_1 + \sum_{k\in B_1'\backslash\{1\}} \pi_{ik}$ with respect to t_{ir} and t_{inr} respectively; $t_{ir}\left(b_{ir}', b_{inr}'\right)$ and $t_{inr}\left(b_{ir}', b_{inr}'\right)$ are the optimal tariffs of the size $b_{ir}' + b_{inr}'$ TA on $3 - b_{ir}'$ regional non-members and $3 - b_{inr}$ non-regional non-members respectively given free trade among the $b_{ir}' + b_{inr}'$ members.) It remains to show that, at $\mathbf{t}\left(b_{ir}', b_{inr}'\right)$, the terms Ξ_i and Φ_i are both strictly negative.

At $\mathbf{t}\left(b_{ir}', b_{inr}'\right)$, $x_{11} =, ..., = x_{1b_{1r}+s_{1r}}$, $x_{14} =, ..., x_{1b_{1nr}+s_{1nr}+3}$ (unless $b_{1nr} = s_{1nr} = 0$, in which case $x_{14} =, ..., x_{1b_{1nr}+s_{1nr}+4}$), and $t = 0$. Also,

$$\begin{aligned}
X_1 &= b_{1r}' x_{11} + \left[3 - b_{1r}'\right] x_{1b_{1r}+s_{1r}+1} + b_{1nr}' x_{1b_{1nr}+s_{1nr}+3} + \left[3 - b_{1nr}'\right] x_{1b_{1nr}+s_{1nr}+4}; \\
T_1 &= \left[3 - b_{1r}'\right] t_{1r}\left(b_{1r}', b_{1nr}'\right) + \left[3 - b_{1nr}'\right] t_{1nr}\left(b_{1r}', b_{1nr}'\right).
\end{aligned}$$

Then we have

$$\begin{aligned}
\Xi_1 &= -4x_{11} + 3d + b_{1nr}' x_{1b_{1nr}+s_{1nr}+3} \\
&\quad - \left(3 - b_{1r}'\right)\left(x_{11} + x_{1b_{1r}+s_{1r}+1} - t_{1r}\left(b_{1r}', b_{1nr}'\right)\right) \\
&\quad - \left(3 - b_{1nr}'\right)\left(x_{1b_{1nr}+s_{1nr}+4} - t_{1nr}\left(b_{1r}', b_{1nr}'\right)\right).
\end{aligned}$$

Now, observing that $x_{11} = x_{1rm}$, $x_{1b_{1r}+s_{1r}+1} = x_{1rnm}$, $x_{1b_{1nr}+s_{1nr}+3} = x_{1nrm}$ and $x_{1b_{1nr}+s_{1nr}+4} = x_{1nrnm}$, we can use (8)-(11) to substitute for x_{11}, $x_{1b_{1r}+s_{1r}+1}$, $x_{1b_{1nr}+s_{1nr}+3}$ and $x_{1b_{1nr}+s_{1nr}+4}$, which obtains

$$\Xi_1 = \frac{-14\left(e - c\right) + \left(9 - 9b_{inr} - 11b_{inr}^2 - 5\left(b_{inr} - 3\right) b_{ir} + 6b_{ir}^2\right) d}{10 + b_{ir}\left(5 + 2b_{ir}\right) + b_{inr}\left(4 + b_{ir}\right) + b_{inr}^2}.$$

First note that $\Xi_1 < 0$ for $d = 0$ and that Ξ_1 is increasing in d. It follows that if $\Xi_1 < 0$ at $d = (e-c)/22$ for all $b_{ir} \in \{1,2,3\}$ and $b_{inr} \in \{0,1,2,3\}$ then this must hold for all $d \in [0,(e-c)/22)$. It is straight forward to verify by substitution that $\Xi_1 < 0$ at $d = (e-c)/22$ for all $b_{ir} \in \{1,2,3\}$ and $b_{inr} \in \{0,1,2,3\}$.

Next observe that

$$\begin{aligned}\Phi_1 &= -4x_{1b_{1r}+s_{1r}+3} - 4d + b'_{1r}x_{11} \\ &\quad - (3-b'_{1nr})\left(x_{1b_{1r}+s_{1r}+3} + x_{1b_{1nr}+s_{1nr}+4} - t_{1nr}\left(b'_{1r}, b'_{1nr}\right)\right) \\ &\quad - (3-b'_{1r})\left(x_{1b_{1r}+s_{1r}+1} - t_{1r}\left(b'_{1r}, b'_{1nr}\right)\right).\end{aligned}$$

Adopting the same basic approach used to simplify Ξ_1, we then have

$$\Phi_1 = \frac{\begin{array}{c}-\left(34 - b_{1r}\left(5+2b_{1r}\right) - 4b_{1nr}\left(b_{1r}-3\right)\right)(e-c) \\ +\left(b_{1r}\left(18b_{1r}-13\right) + b_{1nr}\left(1+b_{1r}\right)\left(2b_{1r}-11\right) + 2b_{1nr}^2\left(b_{1r}-7\right)\right)d\end{array}}{2\left(10 + b_{1r}\left(5+2b_{1r}\right) + b_{1nr}\left(4+b_{1r}\right) + b_{1nr}^2\right)}$$

As was the case for Ξ_1, note that $\Phi_1 < 0$ at $d = 0$ and that Φ_1 is increasing in d. As a result, once again, it is possible to verify by substitution that $\Phi_1 < 0$ at $d = (e-c)/22$ for all $b_{1r} \in \{1,2,3\}$ and $b_{1nr} \in \{0,1,2,3\}$ and that therefore it must hold that $\Phi_1 < 0$ for all $d \in [0,(e-c)/22)$ and all $b_{1r} \in \{1,2,3\}$ and $b_{1nr} \in \{0,1,2,3\}$.

Step 2. We can write the second order condition directly as

$$\begin{aligned}\frac{d^2}{d\mathbf{t}^2}\left(\hat{w}_1 + \sum_{k\in B'_1\setminus\{1\}}\pi_{ik}\right) &= \frac{1}{(7t_{1r})^2}\left(-\left(3-b'_{1r}\right)\left(35 + 15b'_{1r} + 2\left(b'_{1r}\right)^2\right)\Delta t_{1r}^2\right. \\ &\quad + \left(3-b'_{1r}\right)\left(15 + 4b'_{1r} + 2b'_{1nr}\right)\left(s_{1r}t_{1r} + s_{1nr}t_{1nr}\right)\Delta t_{1r} \\ &\quad - \left(3-b'_{1nr}\right)\left(-14 + 15b'_{1nr} + 2\left(b'_{1nr}\right)^2\right)\Delta t_{1nr}^2 \\ &\quad + \left(3-b'_{1nr}\right)\left(15 + 4b'_{1nr} + 2b'_{1r}\right)\left(s_{1r}t_{1r} + s_{1nr}t_{1nr}\right)\Delta t_{1nr} \\ &\quad -2\left(3-b'_{1r}-b'_{1nr}\right)\left(s_{1r}t'_{1r} + s_{1nr}t'_{1nr}\right)^2 \\ &\quad - \left(3-b'_{1nr}\right)\left(7 + 8b'_{1r} - 2b'_{1nr}\left(3-b'_{1r}\right) + 2\left(b'_{1r}\right)^2\right)\Delta t_{1r}\Delta t_{1nr}.\end{aligned}$$

Using the functions for $t_{ir}\left(b_{ir}, b_{inr}\right)$, $t_{ir}\left(b'_{ir}, b'_{inr}\right)$, $t_{inr}\left(b_{ir}, b_{inr}\right)$ and $t_{inr}\left(b'_{ir}, b'_{inr}\right)$, substitution reveals that the second order condition is negative for all feasible values $b'_{1r} \in \{1,2,3\}$ and $b'_{1nr} \in \{0,1,2,3\}$, given $d \in [0,(e-c)/22)$.

Proof of Proposition 5. Without loss of generality, write down two welfare functions for Country 1: $w_1\{\{1,2,3\},\{4,5,6\}\}$ and $w_1\{\{1,4,5,6\},\{2,3\}\}$. (The cases for all other countries are analogous.) The first measures the welfare of Country 1 when it is in a regional

TA with the other countries in its region and all countries in the other region are in a second regional TA. The second measures welfare when Country 1 joins a TA with the countries in the other region while Countries 2 and 3 form a TA. To calculate $w_1\left\{\{1,2,3\},\{4,5,6\}\right\}$, note that Country 1 sets tariffs $t^*_{inr}(3,0)$ on all imports from the other region, and Country 1's exports also face $t^*_{inr}(3,0)$ from all countries in the other region. Trade within regions is free. Using these tariffs in (8) and (11), and substituting appropriately into (7), we obtain

$$w_1\left\{\{1,2,3\},\{4,5,6\}\right\} = \frac{3\left(387\,(e-c)^2 - 134\,(e-c)\,d + 1072d^2\right)}{2450}$$

For $w_1\left\{\{1,4,5,6\},\{2,3\}\right\}$, Country 1 sets $t^*_{ir}(1,3)$ on imports from non-members in its own region. Country 1's exports face tariffs $t^*_{ir}(2,0)$ from non-members in its own region. Trade between Country 1 and the countries in the other region is free. Making appropriate substitutions into (8), (9) and (10), , and substituting appropriately into (7), we obtain

$$w_1\left\{\{1,4,5,6\},\{2,3\}\right\} = \frac{3\left(26442\,(e-c)^2 - 44336\,(e-c)\,d + 92225d^2\right)}{163592}.$$

We can now see that

$$\begin{aligned} w_1\left\{\{1,4,5,6\},\{2,3\}\right\} &> w_1\left\{\{1,2,3\},\{4,5,6\}\right\} \text{ for } d=0; \\ w_1\left\{\{1,4,5,6\},\{2,3\}\right\} &< w_1\left\{\{1,2,3\},\{4,5,6\}\right\} \text{ for } d=(e-c)/22. \end{aligned}$$

We can also see that both $w_1\left\{\{1,2,3\},\{4,5,6\}\right\}$ and $w_1\left\{\{1,4,5,6\},\{2,3\}\right\}$ are decreasing in d for $d \in (0,(e-c)/22)$ but $w_1\left\{\{1,4,5,6\},\{2,3\}\right\}$ is decreasing more rapidly. So we can find a unique value of $d \in (0,(e-c)/22)$, called d', at which $w_1\left\{\{1,2,3\},\{4,5,6\}\right\} = w_1\left\{\{1,4,5,6\},\{2,3\}\right\}$;

$$d' = \frac{3\left(7225156\,(e-c) - 385\sqrt{338226178}\right)\big)}{25290313}(e-c) \simeq 0.017\,(e-c)$$

□

Proof of Proposition 6. By Proposition 3, member welfare of a given TA is decreasing in the size of each of the other TAs that exist. Therefore, the highest feasible level of welfare is achieved when a country is a member of a TA and all non-members of its TA are singletons.

It remains to establish the TA structure that maximizes member welfare (given that all non-members are singletons). First, by Proposition 2, for a singleton to form a two-member TA must increase welfare. We now establish that if both members are in the same region this yields a higher level of welfare than if each member is in a different region. Without loss of generality, assume that Country 1 forms a 2-country TA either with Country 2 in its own region or with Country 4 in the other region. Welfare is $w_1\{\{1,2\},\{3\},\{4\},\{5\},\{6\}\}$ or $w_1\{\{1,4\},\{2\},\{4\},\{5\},\{6\}\}$ respectively. To calculate $w_1\{\{1,2\},\{3\},\{4\},\{5\},\{6\}\}$, note that Country 1 levies a tariff $t^*_{ir}(2,0)$ and $t^*_{inr}(2,0)$ on imports from regional and non-regional non-members respectively. The non-member from R_1 levies a tariff $t^*_{ir}(1,0)$ on imports from Country 1, and non-members from R_2 levy a tariff $t^*_{inr}(1,0)$ on imports from Country 1. Substituting these tariffs into (8)-(11) and substituting appropriately into (7) yields

$$w_1\{\{1,2\},\{3\},\{4\},\{5\},\{6\}\} = \frac{889(e-c)^2 - 999(e-c)d + 2205d^2}{1859}.$$

To calculate $w_1\{\{1,4\},\{2\},\{4\},\{5\},\{6\}\}$, note that Country 1 levies a tariff $t^*_{ir}(1,1)$ and $t^*_{inr}(1,1)$ on imports from regional and non-regional non-members respectively. The non-members from R_1 levy $t^*_{ir}(1,0)$ on imports from Country 1, and non-members from R_2 levy $t^*_{inr}(1,0)$ on imports from Country 1. Substituting these tariffs into (8)-(11) and substituting appropriately into (7) yields

$$w_1\{\{1,4\},\{2\},\{4\},\{5\},\{6\}\} = \frac{7112(e-c)^2 - 4404(e-c)d + 16431d^2}{14872}.$$

Welfare under the two TA configurations is equal for $d = 0$ and the latter yields a lower level of welfare for $d > 0$, with the difference increasing in the size of d.

The same basic approach can be used to establish that the 3-member TA that maximizes a member's welfare is where all members are in the same region, and that a 3-member regional TA yields a higher level of per-member welfare than a 2-member regional TA:

$$w_1(\{1,2,3\},\{4\},\{5\},\{6\}) = \frac{5787(e-c)^2 - 3114(e-c)d + 13362d^2}{11830}.$$

We can also calculate the level of welfare of Country 1 if a non-regional member is included; $w_1\{\{1,2,3,4\},\{5\},\{6\}\}$. In that case, Country 1 imposes a tariff $t^*_{inr}(3,1)$ on imports from non-member, and non-members impose a tariff $t^*_{inr}(1,0)$ on imports from Country 1. Substituting these tariffs into (8), (10) and (11), and making the appropriate substitution into (7), we have

$$w_1\{\{1,2,3,4\},\{5\},\{6\}\} = \frac{333(e-c)^2 - 262(e-c)d + 915d^2}{676}.$$

We can now see that

$$\begin{aligned} w_1\{\{1,2,3,4\},\{5\},\{6\}\} &> w_1\{\{1,2,3\},\{4\},\{5\},\{6\}\} \text{ for } d = 0 \\ w_1\{\{1,2,3,4\},\{5\},\{6\}\} &< w_1\{\{1,2,3\},\{4\},\{5\},\{6\}\} \text{ for } d = (e-c)/22. \end{aligned}$$

We can also see that $w_1\{\{1,2,3,4\},\{5,6\}\}$ is declining in d for $d \in (0,(e-c)/5)$. So we can find a unique value of $d \in (0,(e-c)/22)$, called d'', at which $w_1\{\{1,2,3\},\{4,5,6\}\} = w_1\{\{1,2,3,4\},\{5,6\}\}$;

$$d'' = \frac{1471 - 2\sqrt{433615}}{5301}(e-c) \simeq 0.029\,(e-c)\,.$$

Finally, we must check that a 5-member TA does not yield a higher level of welfare than either a 4-member TA or a 3-member TA. As for all previous cases, a member obtains a higher payoff if all countries in its own region are members of the TA. Thus

$$w_1\{\{1,2,3,4,5\},\{6\}\} = \frac{12145\,(e-c)^2 - 11262\,(e-c)\,d + 37450d^2}{24674}.$$

Since $w_1\{\{1,2,3,4,5\},\{6\}\} < w_1\{\{1,2,3,4\},\{5,6\}\}$ for $d=0$, and since $w_1\{\{1,2,3,4,5\},\{6\}\}$ has a steeper negative slope than $w_1\{\{1,2,3,4\},\{5,6\}\}$, it follows that either $w_1\{\{1,2,3,4\},\{5,6\}\} > w_1\{\{1,2,3,4,5\},\{6\}\}$ or $w_1\{\{1,2,3\},\{4,5,6\}\} > w_1\{\{1,2,3,4,5\},\{6\}\}$. Similar calculations show that free trade yields a lower level of per-member welfare than $w_1\{\{1,2,3,4\},\{5,6\}\}$ or $w_1\{\{1,2,3\},\{4,5,6\}\}$.□

Proof of Proposition 7. There are two cases to consider; $d \in (0,d'')$ and $d \in (d'',(e-c)/22)$. Take $d \in (0,d'')$ first. Given that all countries are singletons in the previous period, by Proposition 5, each country i writes down a strategy s_i listing itself, the two other countries in its region, and one country from the other region. Thus, each country in R_1 names every other country in R_1 in its strategy plus one country from R_2. Symmetrically, again by Proposition 5, each country in R_2 names every other country in R_2 in its strategy plus one country from R_1. But no country in R_1 names every country in R_2 and no country in R_2 names every country in R_1. Therefore, the intersecting set of countries formed by the strategies of countries in R_1 is R_1 itself. So we have a regional TA, $B_1 = R_1$. Symmetrically, the intersecting set of countries formed by the strategies of countries in R_2 is R_2 itself. So we have a second regional TA, $B_2 = R_2$.

Now take $d \in (d'',(e-c)/22)$. The outcome is the same. Given that all countries are singletons in the previous period, by Proposition 5, each country i writes down a strategy s_i listing itself and the two other countries in its region. Thus, each country in R_1 names every other country in R_1 in its strategy. Symmetrically, again by Proposition 5, each country in R_2 names every other country in R_2 in its strategy. Therefore, the intersecting set of countries formed by the strategies of countries in R_1 is R_1 itself. So we have a regional

TA, $B_1 = R_1$. Symmetrically, the intersecting set of countries formed by the strategies of countries in R_2 is R_2 itself. So we have a second regional TA, $B_2 = R_2$.□

Proof of Proposition 8. By Proposition 5, aggregate member welfare increases when a TA expands from 3 members to 6 members (free trade). The two regional TAs, $B_1 = R_1$ and $B_2 = R_2$ are symmetrical, so each country has the same welfare. Thus, the welfare of every country must be increased by the merging of the two TAs to the grand 6 member coalition. Moreover, no country can gain by deviation because a veto of the grand coalition must result in a return to the TA structure of two regional TAs. □

References

[1] Aghion, P., P. Antras and E. Helpman (2004) "Negotiating Free Trade." NBER Working Paper no. 10721, 42 pages.

[2] Arnold, T. and M. Wooders (2005); "Dynamic Club Formation with Coordination." Vanderbilt University typescript.

[3] Bagwell, K. and R. Staiger (1997a); "Multilateral Tariff Cooperation During the Formation of Regional Free Trade Areas." *International Economic Review*, 38: 291-319.

[4] Bagwell, K. and R. Staiger (1997b); "Multilateral Tariff Cooperation During the Formation of Customs Unions." *Journal of International Economics*, 42: 91-123.

[5] Baier, S.L. and J.H. Bergstrand (2004); "Economic Determinants of Free Trade Agreements." *Journal of International Economics*, 64: 29-63.

[6] Bhagwati, J. and A. Panagariya (1996); "Preferential Trading Areas and Multilateralism - Strangers, Friends or Foes?" Published in *The Economics of Preferential Trading Arrangements*, J. Bhagwati, and A. Panagariya (eds): 1-78, American Enterprise Institute, Washington D.C.

[7] Bloch, F. (1996); "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division." *Games and Economic Behavior*, 14: 90-123.

[8] Bloch, F. (2003); "Non-cooperative Models of Coalition Formation in Games with Spillovers." Chapter 2 in C. Carraro (ed) 2003, *The Endogenous Formation of Economic Coalitions* published by Edward Elgar, Cheltenham, UK and Northampton, USA.

[9] Bond, E. (2001); Multilateralism vs. Regionalism: Tariff Cooperation and Interregional Transport Costs." *Regionalism and Multilateralism*, Sajal Lahiri, ed. Routledge Press, 2001.

[10] Bond, R., R. Riezman and C. Syropoulos (2004); "A Strategic and Welfare Theoretic Analysis of Free Trade Areas." *Journal of International Economics*, 64: 1-27.

[11] Bond, E. and C. Syropoulos (1996); "Trading Blocs and the Sustainability of Inter-Regional Cooperation." Published in M. Conzoneri, W.J. Ethier and V. Grilli (eds.) *The New Transatlantic Economy*, London: Cambridge University Press.

[12] Bond, E., C. Syropoulos and L.A. Winters (2001); "Deepening of Regional Integration and Multilateral Trade Agreements." *Journal of International Economics*, 53: 335-361.

[13] Brander, J. and B. Spencer (1984) "Tariff protection and imperfect competition." chapter in H. Kierkowski (ed.) *Monopolistic Competition and International Trade*, Oxford University Press, Oxford.

[14] Frankel, J., E. Stein and S. Wei (1995); "Trading Blocs and the Americas: The Natural, the Unnatural and the Super-natural." *Journal of Development Economics*, 47: 61-95.

[15] Goyal, S. and S. Joshi (2005); "Bilateralism and Free Trade." *International Economic Review* forthcoming.

[16] Hart, S. and M. Kurz (1983); "Endogenous Formation of Coalitions." *Econometrica*, 51(4): 1047-1064.

[17] Kennan, J. and R. Riezman (1990); "Optimal Tariff Equilibria with Customs Unions." *Canadian Journal of Economics*, 23: 70-83.

[18] Krishna, P. "Regionalism and Multilateralism: A Political Economy Approach." *Quarterly Journal of Economics*, 113(1): 227-251.

[19] Krugman, P. (1991); "The Move to Free Trade Zones." Published in *Policy Implications of Trade and Currency Zones*, Federal Reserve Bank of Kansas City.

[20] Panagariya, A. (1998) "Do Transport Costs Justify Regional Preferential Trading Arrangements? No." *Weltwirtschaftliches Archiv/Review of World Economics* 134(2): 280-301.

[21] Riezman (1985); "Customs Unions and the Core." *Journal of International Economics*, 19: 355-66.

[22] Riezman, R. (1999); "Can Bilateral Trade Agreements Help to Induce Free Trade?" *Canadian Journal of Economics*, 32(3): 751-766.

[23] Seidmann, D.J. and E. Winter (1998); "A Theory of Gradual Coalition Formation." *Review of Economic Studies*, 65(4): 793-815.

[24] Summers, L. (1991) "Regionalism and the World Trading System." Published in *Policy Implications of Trade and Currency Zones*, Federal Reserve Bank of Kansas City.

[25] Wonnacott, P., and M. Lutz (1987); "Is There a Case for Free Trade Areas?" In Jeffrey Schott (ed.) *Free Trade Areas and US Trade Policy*, Institute for International Economics, Washington DC.

[26] WTO (2000); "Mapping of Regional Trade Agreements" World Trade Organization, Geneva. Available online at http://www.wto.org.

[27] Yi, S.-S. (1996); "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good" *Journal of International Economics*, 41: 151-175.

[28] Yi, S.-S. (2003); "Endogenous Formation of Economic Coalitions: A Survey of the Partition Function Approach." Chapter 2 in C. Carraro (ed) 2003, *The Endogenous Formation of Economic Coalitions* published by Edward Elgar, Cheltemham, UK and Northampton, USA.

[29] Zissimos, B. (2002); "Why Are Trade Agreements Regional: A Theory based on Noncooperative Networks." Warwick Economic Research Paper no 652.