

# Discussion Papers in Economics

## Endogenous Trading Bloc Formation

### In a North-South Global Economy

**Satya P.Das**

and

**Subhadip Ghosh**

**Discussion Paper 03-04**

**August 2003**



Indian Statistical Institute, Delhi  
Planning Unit

7 S.J.S. Sansanwal Marg, New Delhi 110 016, India

# Endogenous Trading Bloc Formation in a North-South Global Economy

Satya P. Das, Indian Statistical Institute

Subhadip Ghosh, Indian Statistical Institute\*

August 7, 2003

## Abstract

Majority of the trading blocs to date are between similar countries, rather than between developed and developing countries. This paper provides a rationale for why trading blocs among similar countries may arise as an equilibrium phenomenon. It develops a model of an asymmetric world economy, in which there are at least four countries. The countries are differentiated with respect to their market size and market structure.

JEL Classifications: F02, F13, F15.

---

\*Corresponding author: Indian Statistical Institute, 7 S.J.S. Sansanwal Marg, New Delhi: 110016, India. E-mail: subhadip@isid.ac.in. Tel: 91-11-26564990. Fax: 91-11-26856779. The authors thank Pravin Krishna and Anindya Bhattacharya for valuable comments. Comments from participants in the Conference on International Trade and Factor Mobility held at Kansai University, Japan, and International Workshop on Economic Theory held in Center for Studies in Social Sciences, Kolkata, India are also appreciated.

# 1 Introduction

There has been a spectacular rise in the number of Preferential Trading Arrangements (PTAs) since the 1980s: compared to 65 PTAs notified to the GATT/WTO up to 1980, at the end of 1999 there were 194 notified PTAs – that is, the number of PTAs nearly trebled in just 20 years (World Bank (2000)). These PTAs now account for well over half of the world trade. Which type of countries typically form a trading bloc?

World Bank (2000, page ix) notes that regional integration agreements have taken place “among high-income countries, among low-income countries, and, more recently, starting with North American Free Trade Area (NAFTA) – between high-income and developing countries.” This is indicative that, generally speaking, most PTAs have been formed among similar countries, while “North-South” PTAs, of more recent origin, are rather very few in number. Bhagwati and Panagariya (1996) provide a comprehensive list of PTAs notified to GATT and WTO up to the year 1995. From this, one can calculate that out of 134 PTAs notified in the GATT and WTO up to June 1, 1995 only about 15 can be regarded as of the North-South kind; the rest are either North-North or South-South. As Ray (1998, page 730) notes, “.. the vast majority of regional agreements have taken place among countries at relatively *similar* levels of economic development.”

Hence, to ‘stylize’ this fact, PTAs are formed among similar countries. Of course, there are prominent exceptions like NAFTA in which a developing country (Mexico) is a full partner with rich countries (U.S. and Canada).<sup>1</sup>

Indeed, such a pattern is not so apparent, because the standard trade theory based on comparative advantage says that countries benefit from freer trade by exploiting their differences. It is noteworthy that, for these reasons, Ray (1998), on the same page, states that “the paucity of North-South arrangements is paradoxical.”

There are two objectives of this paper. One is to provide a model that ‘generates’ PTAs among similar countries. More specifically, it considers endogenous trading bloc formation in which the equilibrium outcome is PTAs among similar, rather than among dissimilar countries. In what follows, we refer to this case as ‘polarization’. The second objective is to understand which type of countries want to block global free trade and under which conditions.

On the theoretical front, a huge literature exists on the economics of PTAs, a recent survey of which is contained in Panagariya (2000). Starting from Viner, Lipsey and Meade, to more recent ones like Levy (1997), Krishna (1998) and Bond, Syropoulos and Winters (2001) among many others, an ‘abstract’ three-country framework has remained as the core model of analysis. Other recent papers such as Krugman (1991), Yi (1996, 2000) and Freund (2000), on the other hand simply assume that all trading

---

<sup>1</sup>Inclusion of Portugal, Greece and Spain into EU can also be considered similar to NAFTA, although the difference between these countries and other EU members is not as stark as that between US-Canada and Mexico.

countries are similar. In contrast, this paper develops a novel *four-country* model in an *asymmetric* world economy, in which a country has a *choice* between forming a bloc with another similar country or with a dissimilar country. The merit of having a 4-country setup, as opposed to the standard 3-country model, is that it can accommodate the division of world economy into more than one trade bloc at a time whereas in an asymmetric 3-country setup, the scope of combining with a similar country cannot be available for all three countries.<sup>2</sup>

Among existing works, Ethier (1998) comes closest to our analysis. Yet it is very different. It assumes that a North-North PTA is formed initially. Given the presence of international scale economies that are external to the firm, this PTA lowers the cost of production of an intermediate input and makes the production of this input feasible in the South, which hitherto was not possible. This leads to the formation of North-South trading blocs, where Northern countries outsource a part of its intermediate input production to the partner South country, and in return get exclusive right to sell a final good in the South. In contrast, in this paper all countries move simultaneously and focus is on whether the equilibrium outcome, if not global free trade, is polarization or North-South mixing.

In general, there are of course many sources of dissimilarity among countries. The primary focus of this paper is on difference in market size: some countries have bigger markets than others. This is meant to capture that the magnitude of market access to one another is an important factor for a country in its consideration to form FTA with some other countries. In terms of a North-South world economy, we may think the North countries as the high-demand countries and South countries as the low-demand countries.<sup>3</sup> Within this framework, we also consider another source of dissimilarity, namely, that in market structure.

Evidently a standard, competitive, model of international trade will not “deliver” the stylized fact. In what follows, we use an imperfectly competitive model as in Krishna (1998), which has oligopoly competition and market segmentation. This may raise an immediate reservation that such a framework characterizes trade among developed countries rather than between developed and developing countries. However, in recent decades many manufacturing sectors in the world economy are dominated by firms from both developed and developing countries. This is especially true if we include East Asian countries like Korea and Taiwan in the latter category.<sup>4</sup> In the automobiles and electronics industries, for example, firms like Daewoo, Hyundai and Samsung have significant market shares in both developed and developing countries. Besides these countries, others like India, Pakistan and China are now gaining increasing market shares in many light manufacturing industries like garments, kitchenware and

---

<sup>2</sup>Later, our central result on equilibrium PTAs is shown to be not dependent on the number of countries being equal to four.

<sup>3</sup>Even when one thinks of very populous developing countries like China or India, their per-capita income adjusted market size is smaller than, say, the U.S. or the EU.

<sup>4</sup>Recall that in our model, developed and developing countries are distinguished by the difference in their market sizes: developing countries are those whose market sizes are smaller. These East Asian countries all have relatively small domestic markets and therefore would fit as developing countries in our model economy.

mechanized toys. Moreover, many of these industries are oligopolistic. Thus an oligopoly framework does characterize competition among firms from developed and developing countries – at least to a significant extent.

However, note that such a framework is not built on symmetry only. As mentioned earlier, countries may be dissimilar with respect to market size and market structure. In other words, such a framework accommodates both inter-industry and intra-industry trade.<sup>5</sup> Our model is thus a hybrid one, having elements of gains from trade among similar countries as well as comparative advantage.

In terms of “dynamics” of forming trading blocs, our analysis assumes that countries, initially, are in a state of unilateralism, when they plan to form trading blocs or to move to global free trade directly. The latter option is nothing but multilateralism. We assume that trading blocs are in the form of FTAs and there are no side payments.<sup>6</sup>

Our basic model assumes that direct negotiations on global free trade (from a position of unilateralism) takes one time period, whereas negotiating an FTA is instantaneous. This assumption intends to capture that the transaction costs of negotiating and implementing multilateral talks are greater than those of negotiating FTAs. This is accommodated by considering a two-period model. Later, this assumption is relaxed, and both FTAs and global free trade are assumed to be negotiable instantaneously.

Endogeneity in the formation of trading blocs in an asymmetric world economy is the central feature of our model. It is different from Yi (1996, 2000), who considers a world economy with symmetric countries. In our model whether or not to form a PTA *and* the identity of PTA partner(s) are both endogenous. Further, Yi uses Nash equilibrium and Sub-Game-Perfect-Nash equilibrium as the equilibrium concepts. But, as noted in Sengupta et al (1997), the equilibrium concept in a model considering such endogenous bloc formation must allow joint deviation by a group of countries, and as such Nash equilibrium – which takes into account only unilateral deviations – is not appropriate. Sengupta et al suggest Coalition-Proof-Nash-Equilibrium (CPNE) as the suitable refinement. Since one variant of our model considers a multi-stage game, we do not use CPNE. Instead, we use Perfectly Coalition-Proof-Nash-Equilibrium (Peleg (1992))– which induces a CPNE in every subgame. Further, CPNE permits only internal deviations – that is, only deviations by coalitions which are subsets of coalitions in the existing coalition structure. This model however puts no such restrictions on deviations by groups of countries. As such the equilibrium we consider is in some way a more refined version of Perfectly Coalition-Proof-Nash-Equilibrium.

A general result that comes out of our model is that the world economy evolves into either global free

---

<sup>5</sup>This is consistent with the evidence that, between developed and developing countries, not only the volume of inter-industry trade has grown in recent decades, so has intra-industry trade. For instance, Nilsson (1997) reports increasing trend in intra-industry trade between EU and developing countries in general. Kandogan (2003) notes the same between EU and CEEC (Central and East European) countries. According to Clark and Stanley (1999), such trade is about 50% of all trade in manufacturing between the United States and the developing countries. Thus both inter-industry and intra-industry trades are important.

<sup>6</sup>To date, among PTAs, FTAs are more common than customs unions.

trade, or into FTA blocs that are polarized, i.e. equal-market-size countries match with one another. In this sense it is consistent with the stylized fact about which type of countries form FTAs. What happens is that the high-demand countries want to and hence match with each other in equilibrium. A low-demand country's most preferred choice is also to be a partner of a high-demand country. But it is not reciprocated. Consequently, low-demand countries match with each other residually – and are still better off compared to unilateralism. Moreover, from a situation of polarization where there are no standard gains from trade among similar countries to be realized but potential gains from trading with dissimilar countries are present, there is still no profitable deviation except possibly to the grand global free trade. This of course means that potential costs outweigh potential gains.

As mentioned earlier, one of our objectives is to also understand which types of countries may object to global free trade and why. In various cases in terms of market structure differences and the number of countries, it is found that in most cases, if at all global free trade is blocked, it is done so by the high-demand countries. Only in one case, the low-demand countries are the blockers.<sup>7</sup>

The paper is organized as follows. In section 2, the rules of the game and the equilibrium coalition formation concept are outlined in the context of a two-period world economy. Section 3 examines our core 4-country model. Section 4 considers a single-period static world economy in which PTAs and global free trade have the same time frame. Difference in market structure is introduced and the model is extended beyond four countries. Section 5 offers some concluding remarks.

## 2 Rules of the Game

In the beginning of period I, each country is practising unilateralism, i.e. it is following its own trade policy independently. Starting from this situation, countries may opt for multilateralism, or FTAs or they may continue to remain unilateral. Multilateral negotiation, if it occurs, takes the whole of period I and hence countries remain unilateral in period I, and the global free trade regime is operational in the beginning of period II. The world ends after period II.

FTA negotiations, on the other hand, are assumed to be completed instantaneously. Hence, if formed, they are operational from the beginning of period I itself. Such negotiations may be undertaken again in the beginning of period II. Countries within an FTA, or a country which has remained “isolated” at the end of period I, may join another country in isolation or an FTA.<sup>8</sup> The outcome in period II may or may not be global free trade.

---

<sup>7</sup>It is well-known that developing countries have in the past voiced their opposition to global free trade – probably more so than have the developed countries. However, their “fear” has been typically founded on market power of firms and governments from the developed countries. Our model, however, is not equipped to handle these concerns. But it is able to uncover some of the reasons as why developed countries may be opposed to global free trade.

<sup>8</sup>We allow for the possibility of a country in an FTA to break away from that FTA and be a part of some other FTA.

## Formal Representation and Equilibrium Concept

The multi-stage game outlined above is one with perfect information. Following general representations of these sort of games by Fudenberg and Levine (1983) and Maskin and Tirole (2001), define a game  $G$  with  $m$  players (countries  $A, B, C, D, \dots$ ) and two periods (indexed by  $t=I, II$ ). In every period  $t$ , a player (any country)  $i$  chooses an action  $\mathcal{A}_t^i(h_t)$  where  $\mathcal{A}_t^i$  may depend on actions chosen in earlier periods through history  $h_t$ . By  $h_t$  we mean the sequence of actions chosen before period  $t$ , *i.e.*  $h_t = (\mathcal{A}_1, \dots, \mathcal{A}_{t-1})$  where  $\mathcal{A}_t = (\mathcal{A}_t^A, \mathcal{A}_t^B, \mathcal{A}_t^C, \mathcal{A}_t^D, \dots)$ . The history at any period  $t$  is a common knowledge.

In the beginning of period I, a player's action set consists of three actions: it can vote for or opt (a) multilateralism, (b) regionalism, that is, name an FTA partner or (c) stay unilateral. Multilateralism will be implemented if and only if all countries vote for it – *i.e.* by consensus. Regionalism, on the other hand, requires only a pair wise matching. A particular country will stay unilateral in period I if (i) multilateralism is voted and it is the consensus, since it takes one time period for multilateral negotiations to be completed, (ii) it votes for multilateralism but it is not a consensus, (iii) it names a FTA partner, but it is not reciprocated or (iv) it opts for unilateralism.

In period II, if all countries have voted for multilateralism, the action space is a null set. Otherwise, the action space again consists of (a) naming an (already formed) FTA, or any other single country as its FTA partner, or (b) continuing to stay unilateral. It is assumed that one country cannot belong to two different trading blocs.<sup>9</sup> To this extent the action space is constrained by history. The total payoff from the game for any country is the discounted sum of its welfare from the two periods.

Being a multi-period game, the most common solution concept that can be used is the Sub-Game-Perfect-Nash-equilibrium (SPNE), which induces a Nash equilibrium in every subgame of the multi-stage game. However, Nash equilibrium considers only unilateral deviations. A model of coalition formation such as this must instead allow for a group of countries to deviate jointly, if that benefits all the deviating countries. As such, we should consider a refinement of Nash equilibrium that ensures that equilibrium profiles are immune to individual as well as joint deviations.

We find that the most suitable refinement for our model is the Perfectly Coalition-Proof-Nash-equilibrium or PCPNE, a concept introduced by Peleg (1992). A strategy is a Coalition-Proof-Nash-equilibrium (CPNE) if no coalition (set of countries) can make a deviation that is both profitable to each of its member countries, and also cannot be disrupted by further deviations by subsets of the deviating coalition. A strategy is PCPNE in an extensive form game if it induces a CPNE to any subgame of the game. However, one limitation of PCPNE is that it considers only internal deviations from an existing coalition structure, that is, a country within a coalition (PTA) cannot forge an agreement with any member of a complementary coalition. We however allow for the countries in two different existing

---

<sup>9</sup>It means that, if countries  $A$  and  $B$  have formed an FTA, then it cannot be the case that later  $A$  – but not  $B$  – forms an FTA with country  $C$ , unless the FTA negotiation between  $A$  and  $B$  are broken.

coalitions to coordinate on a joint deviation, if such a deviation makes all the deviating countries better off. Thus, the equilibrium concept used here is a modified and more refined version of PCPNE. However, with slight abuse of definitions, for the sake of brevity, throughout the paper we will use the term PCPNE instead of a modified PCPNE.

### 3 The basic 4-country model

There are four countries,  $A$ ,  $B$ ,  $C$  and  $D$ . Countries  $A$  and  $B$  are identical, so are  $C$  and  $D$ . The only difference between  $(A, B)$  and  $(C, D)$  lies in their market sizes:  $A$  and  $B$  have a bigger market size than do  $C$  and  $D$ . In terms of North-South paradigm,  $A$  and  $B$  can be thought of as North countries while  $C$  and  $D$  are the South countries.

Each country produces two goods –  $Z$  and  $M$ . In the supply side, there is only one factor of production, labor, which is used in both sectors. There are constant-returns, which imply that the labor coefficients are given. Good  $M$  is competitively produced (along with free entry and exit) so that the wage rate is given. The market for good  $Z$  is assumed to be oligopolistic. Given the labor coefficient and the wage rate, the marginal cost of producing good  $Z$  in terms of good  $M$  (taken as the numeraire) is a constant. Let it be defined as  $c$ .

Following Krishna (1998), the markets are assumed to be segmented, so that firms compete in each country separately. As discussed in the Introduction, this framework captures inter-industry and intra-industry trade, both of which are quantitatively important in the total trade in manufacturing between developed and developing countries. Analytically, market segmentation yields that profits across markets are independent of each other. Hence any particular country's optimal tariff is independent of that by some other country. This considerably simplifies welfare comparisons across various regimes. We also assume quantity competition, as in Yi (1996, 2000), Krishna (1998) and Freund (2000). For simplicity, let there be one firm from each country. (This assumption is relaxed later.)

The utility function of a representative consumer in any high-demand country, say  $A$ , and in any low-demand country, say  $C$ , are respectively:

$$G^A(z, m) = \alpha_H z^A - \frac{1}{2}(z^A)^2 + m^A; \quad G^C(z, m) = \alpha_L z^C - \frac{1}{2}(z^C)^2 + m^C, \quad (1)$$

where  $z^A$  ( $z^C$ ) is the total consumption of good  $Z$  in country  $A$  ( $C$ ), and  $m^A$  ( $m^C$ ) denotes the same for the numeraire good  $M$ . The assumption of quasilinear utility functions implies a partial equilibrium approach, looking at the market for good  $Z$  only. Let  $\alpha_H > \alpha_L$ . This captures that  $(A, B)$  and  $(C, D)$  are respectively high-demand and low-demand countries.<sup>10</sup>

---

<sup>10</sup>Given  $\alpha_H > \alpha_L$ ,  $(A, B)$  and  $(C, D)$  can also be interpreted as high-*income* (North) and low-*income* (South) countries. Let the utility function be  $G(z, m) = \alpha(m)z - \frac{1}{2}(z)^2 + m$ , where  $\alpha(m)$  is an increasing step function of  $M$ , the amount



It is assumed that the tariff revenues are transferred back to the consumers. A country  $k$ 's welfare,  $W^k$ , has four components: the domestic consumers surplus ( $CS^k$ ), the home firm's profit ( $\pi_k^k$ ) in its home market as well as its profits in the export markets ( $\pi_k^j$ ,  $j \neq k$ ) and the tariff revenue ( $R^k$ ). Thus

$$W^k = CS^k + \pi_k^k + \sum_{j \neq k} \pi_k^j + R^k, \quad j, k \in \{A, B, C, D\}, k \neq j$$

where superscripts refer to the country where competition is taking place, while the subscripts in a firm's profit expression refer to the country from which the firm originates. This can be re-expressed, for country  $A$  for example, as the following, which is easier to manipulate.

$$W^A = G^A(z^A) - cz^A - (\pi_B^A + \pi_C^A + \pi_D^A) + (\pi_A^B + \pi_A^C + \pi_A^D), \quad (2)$$

### 3.1 Unilateralism: the Status Quo

We now characterize unilateralism, the status quo. Let  $t^k$  be the specific tariff imposed by country  $k$  on its imports from all other countries. Then the effective marginal costs of exporting to this country is  $c^k = c + t^k$ .

Consider first competition in country  $A$ . Introduce the following notations:

$z_A^A$ : the home firm's supply in the  $A$  market.

$P^A$ : price of good  $Z$  in country  $A$ .

$t^A$ : tariff set by the government of country  $A$ .

The objective of the home firm ( $A$ ) and foreign firms ( $j$ ) are respectively

$$\max_{z_A^A} \pi_A^A = (P^A - c)z_A^A, \quad \max_{z_j^A} \pi_j^A = (P^A - c - t^A)z_j^A; \quad j = \{B, C, D\}$$

The unique Cournot-Nash equilibrium solutions in stage 2, along with the corresponding profit expressions, are given by

$$z_A^A = \frac{a_H + 3t^A}{5}, \quad z_A^j = \frac{a_H - 2t^A}{5}, \quad \pi_A^A = \left(\frac{a_H + 3t^A}{5}\right)^2, \quad \pi_A^j = \left(\frac{a_H - 2t^A}{5}\right)^2, \quad j = \{B, C, D\}; \quad (3)$$

where  $a_H \equiv \alpha_H - c$ . Similarly, by defining  $a_L \equiv \alpha_L - c$ , the profit of the  $A$ -firm in other markets will be

$$\pi_A^B = \left(\frac{a_H - 2t^B}{5}\right)^2, \quad \pi_A^C = \left(\frac{a_L - 2t^C}{5}\right)^2, \quad \pi_A^D = \left(\frac{a_L - 2t^D}{5}\right)^2. \quad (4)$$

---

consumed of the numeraire good. Consider two income levels  $I_H$  and  $I_L$  such that  $I_H > I_L$ . If  $|I_H - I_L|$  is large enough, the above utility function imply that  $\alpha_H \equiv \alpha(m)|_{I=I_H} > \alpha_L \equiv \alpha(m)|_{I=I_L}$ . Furthermore, if the length of a step in the step function  $\alpha(m)$  is large enough,  $\alpha_H$  and  $\alpha_L$  can be taken as parameters (with a single value) in face of income and price movements due to trade policy changes.

Substituting (3)-(4) into (2), we have

$$W^A(t^A) = a_H \left( \frac{4a_H - 3t^A}{5} \right) - \frac{1}{2} \left( \frac{4a_H - 3t^A}{5} \right)^2 - 3 \left( \frac{a_H - 2t^A}{5} \right)^2 + \left( \frac{a_H - 2t^B}{5} \right)^2 + \left( \frac{a_L - 2t^C}{5} \right)^2 + \left( \frac{a_L - 2t^D}{5} \right)^2, \quad (5)$$

where the last three terms, which represent the profits of the  $A$ -firm in other countries, are independent of  $t^A$ . This is maximized with respect to  $t^A$ . The optimal tariff is given by:

$$t_U^A = \frac{3a_H}{11}, \quad (6)$$

where  $U$  stands for “unilateral”. Markets being segmented, country  $k$ 's tariff does not depend on country  $j$ 's tariff on its product.

By symmetry, (6) is the optimal tariff for  $B$ . Likewise, a low-demand country's optimal tariff is given by

$$t_U^C = t_U^D = \frac{3a_L}{11}. \quad (7)$$

Substituting (6)-(7) into (5) and doing the same for low-demand countries, the reduced-form welfare expressions under unilateralism are:

$$W_U^A = W_U^B = \frac{101a_H^2 + 4a_L^2}{242}; \quad W_U^C = W_U^D = \frac{101a_L^2 + 4a_H^2}{242}, \quad (8)$$

To ensure that each firm sells positive quantities, we need to assume the following regularity condition:

$$a_L = \alpha_L - c > 0. \quad (R1)$$

From (8), notice that tariffs by the high-demand country are higher. This is because having a larger market offers their governments a greater scope to exercise the standard monopsonistic power.<sup>11</sup>

### 3.2 Coalition formation

Figure 1 depicts the game in extensive form. There are three period-I possibilities, (a) the countries remain unilateral, either because of opting for multilateralism, or choosing to remain unilateral (without any commitment to global free trade in period II), (b) there are two FTAs of two countries each, or (c) there is one FTA of two countries, while two others remain unilateral or (d) there is a three-

---

<sup>11</sup>The implication of this result in a North South perspective may be at odds with what is observed empirically, that is, South countries generally impose higher tariffs than do the North countries. But if South puts sufficiently higher weight on their tariff revenues than does the North (which can be justified because of their greater need for revenues to fund their enormous subsidy expenditures), then North tariffs will be lower than the South tariffs. For our purpose it is the *dynamics* of tariffs as the PTAs are formed which is more important than the ranking of tariffs across North and South at a given point of time.

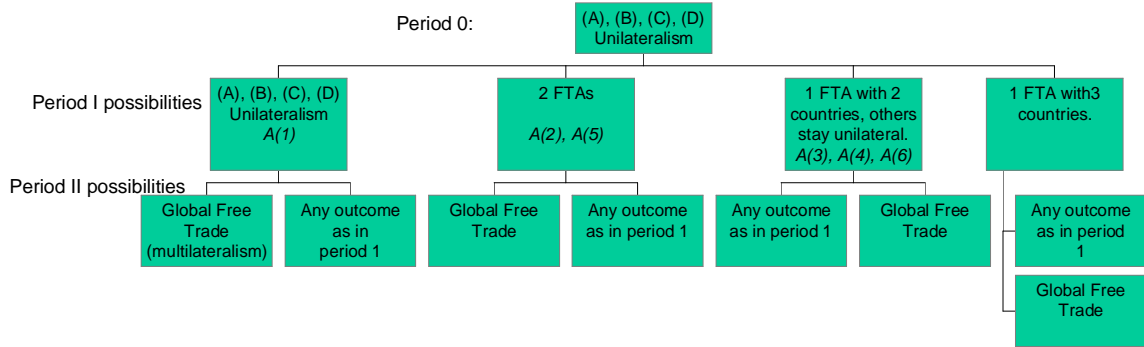


Figure 1: The Game Tree

country FTA, while one country is left out. Corresponding to each of these possible outcomes in period I, the consequent period II possibilities are then depicted. Note that the left-most branch shows multilateralism. The question is, which among these numerous possibilities emerges as the equilibrium coalition structure of our model world economy in period I and period II.

To solve this, we use backward induction.

### Period-II Subgame

We seek CPNE corresponding to every possible coalition (FTA) structure that may already be formed in period I. Note that besides CPNEs there is one other outcome in period II: global free trade via multilateralism. It evolves from unilateralism ( $\mathcal{A}_I(1)$ ) as a multilateral commitment rather than as a negotiated outcome from any coalition formed in period I. Hence, it is not a candidate for CPNE in period II. But it is a candidate for the PCPNE in period I.

Recall that  $\mathcal{A}_t$  denotes the action set of all countries in period  $t$ . For  $t = I$ , further define  $\mathcal{A}_I(j)$  as the action set  $j$  that is actually played, among all feasible action sets in period I. In Table 1, it is shown that  $j \in \{1, 8\}$ . The solution of the period-II subgame constitutes our first proposition.

**Table 1: Period-II Outcomes**

<i>Actions taken in period I: <math>\mathcal{A}_I</math></i>	<i>Corresponding CPNE (except the last entry)</i>
$\mathcal{A}_I(1): \{A, B, C, D\}$	polarization: $\{A-B, C-D\}$
$\mathcal{A}_I(2): \{A-B, C, D\}$	polarization if $\rho > \bar{\rho} \equiv 1.14$ global free trade, i.e., $\{A-B-C-D\}$ if $\rho \leq \bar{\rho}$
$\mathcal{A}_I(3): \{A, B, C-D\}$	same as above
$\mathcal{A}_I(4): \{i-m, k, l\}$ <i>i and k: two high-demand countries</i> <i>m and l: two low-demand countries</i>	same as above
$\mathcal{A}_I(5): \{A-C, B-D\}$ or $\{A-D, B-C\}$	same as above
$\mathcal{A}_I(6): \{i-j-k, m\}$ <i>i, j, k, m: the four countries</i>	same as above
$\mathcal{A}_I(7): \{A-B, C-D\}$	same as above
$\mathcal{A}_I(8):$ multilateralism i.e. $\{A, B, C, D\}$	global free trade by commitment

**Proposition 1** *The period-II CPNEs are as given in the right column of Table 1 (except for the last row), which correspond to the respective action set,  $\mathcal{A}_I(1)$  through  $\mathcal{A}_I(7)$ , indicated in the left column.*

**Proof:** See Appendix 1. ■

It is remarkable that the strategy sets  $\mathcal{A}_I(2)$  through  $\mathcal{A}_I(7)$  in period I lead to the *same* and unique CPNE in period II: polarization if the market size difference is large enough and global free trade otherwise. Moreover, this is same as when  $\mathcal{A}_I(1)$  is undertaken in period I except that global free trade is not feasible by assumption (as a jump to global free trade from unilateralism in a single period is ruled out).

Instead of discussing how the same CPNE emerges from various initial strategy sets individually, it seems useful to understand the general effects and key results that lead to Proposition 1. There are two such effects and five such results.

*Result 1:* Consider a firm's profit in the domestic and a foreign market. If the foreign market size is same as or greater than that of the domestic market size, a symmetric reduction in (specific) tariff increases the firm's profit.

This is because a tariff reduction in the foreign market is equivalent to a decrease in the own marginal cost of a domestic firm, whereas a tariff reduction in the domestic market is equivalent to an increase in the rival firms' marginal cost. Straightforward calculation based profit expressions in Cournot oligopoly equilibrium yields that, *ceteris paribus*, the effect of the former on the domestic firm's profits is higher in magnitude than the effect of the latter. This can be interpreted as a **relative market access effect**.

Assume for now that there is an FTA structure and it is in the form of polarization in period I. With the help of *Result 1*, *Result 2* below explains why the CPNE in period II is polarization or global free trade according as the size difference between the two pairs of countries is large or small.

*Result 2:* To a high-demand country, global free trade yields higher or less *one-period* welfare compared to polarization as  $\rho \lesseqgtr \bar{\rho}$ . However, to a low-demand country global free trade always yields higher one-period welfare.<sup>12</sup>

Starting from polarization, in global free trade the low-demand South countries offer tariff concessions in smaller markets to firms who are equally efficient as its own firms, whereas they receive tariff-free access for their firms in bigger markets. In view of *Result 1*, this is a more profitable proposition. However, for high-demand countries, global free trade offers tariff-free access in smaller markets, while they have to offer tariff concessions in bigger markets. In view of *Result 1*, this is profitable if and only if a South country's market size is not too small compared to that of a North country. Otherwise, global free trade yields less profit to high-demand North countries compared to polarization. Of course, *Result 2* refers to total welfare which, along with profits, includes consumer surplus and tariff revenues. But profit rankings happen to dictate the welfare ranking.

Now suppose that an FTA can consist of two countries only and it can be in the form of either polarization or North-South mixing (e.g. *A-C* and *B-D*). The next result implies why North-South mixing cannot be a CPNE.

*Result 3:* From unilateralism, each country is better off – in terms of one-period welfare – by matching with a same-size country and best off by matching with a high-demand country.<sup>13</sup>

The implication is that North-South mixing is not strategy-proof. The two North countries match with each other as their best option and the two South countries match residually (and are still better off compared to unilateralism). The reason is that, for North countries who have the same market sizes and which are greater than those in South countries the incentive to match with each other follows from the underlying relative market access effect; this is somewhat similar to the usual gains from trade among similar countries. The incentive of a low-demand South country to match with a high-demand North country stems from the market access effect also; however, note that this is *not* same as the standard gains from trade among similar countries.

The following result is a corollary of *Result 3*.

*Result 4:* One-period welfare of each country is higher under polarization than under unilateralism.

The next result shows that any 3-country PTA is not strategy-proof against polarization.

*Result 5:* Starting from a situation of polarization, a (i) 3-country FTA with two North countries and one South country yields less welfare to the North countries and a (ii) 3-country FTA with one North country and two South countries yields less welfare to the deviating North country.

Given polarization, for any enlarged bloc, there are no further usual gains from trade with similar countries. Yet, part (i) says that the North bloc wouldn't benefit from free trade with one (dissimilar)

---

<sup>12</sup>It is proven in Appendix 1.

<sup>13</sup>It is also proven in Appendix 1.

South country.<sup>14</sup> It is because the North bloc would get access only to a smaller market, but give access to an equally efficient South firm in two bigger markets. This relative market access effect yields less profit and welfare to the North bloc.

Consider (ii) now. There are two subcases. Suppose that the combined market size of the South bloc is less than or equal to that of any one North country (i.e.  $2a_L \leq a_H$ ). It is then obvious that the deviating North country wouldn't benefit from this. Suppose  $2a_L > a_H$ . Then the North country would gain access to a larger market, which tends to benefit it. However, by joining the South bloc, the (one) firm in the deviating North country has to compete now in equal terms with *two* equally efficient firms from the two South countries. This is a **relative competition effect**, which tends to lower profits and welfare and dominates the relative market access effect. Hence a North country wouldn't join the South bloc.

*Result 5* implies that polarization is coalition-proof against any kind of 3-country FTA. *Results 2-5* together imply that the CPNE in period II is either global free trade or polarization. The important point to note is that the rationale behind the polarization outcome goes beyond the usual gains from trade among similar countries. The relative market access effect and the relative competition effect explain this outcome.

Next, who, if at all, blocks global free trade? *Result 2* readily implies

**Proposition 2** *South countries do not wish to block global free trade, whereas North countries block it if  $\rho > \bar{\rho}$ .*

We now move on to analyze Period-I game.

## Period-I game

The difference between the period-II subgame and the period-I game is that while in period II the objective of a country is to maximize its welfare in that period only, in period I it is to maximize the discounted sum of the welfare over the two periods. The equilibrium concept used is that of PCPNE, which has been explained earlier. Proposition 3 summarizes the result, which is the full solution of the game.

**Proposition 3** *The PCPNE has a unique solution, given by  $\mathcal{A}_I(7)$ . Thus the world economy is polarized in period I. It is followed, in period II, by either polarization again or global free trade according as  $\rho \gtrless \bar{\rho}$ .*

**Proof:** The task is to show that the action set defined by  $\mathcal{A}_I(7)$  is strategy-proof. First consider  $\mathcal{A}_I(1)$ , in which the countries choose to stay unilateral, and the corresponding CPNE in period II is

---

<sup>14</sup>As will be seen in the next section, this holds even when there is another source of dissimilarity, namely, in terms of market structure.

polarization. In view of the relative market access effect which underlies *Result 4*,  $\mathcal{A}_I(7)$ , relative to  $\mathcal{A}_I(1)$ , yields higher period-I welfare to any country. If  $\rho \geq \bar{\rho}$ , there is no difference in period-II outcome between  $\mathcal{A}_I(7)$  and  $\mathcal{A}_I(1)$  (it is polarization), and, if  $\rho < \bar{\rho}$ , all countries prefer global free trade to polarization. Hence no country is worse off in period II under  $\mathcal{A}_I(7)$ . It follows that  $\mathcal{A}_I(7)$  strongly dominates  $\mathcal{A}_I(1)$  in the sense that all the countries are better off under  $\mathcal{A}_I(7)$  than under  $\mathcal{A}_I(1)$ .

The strategy sets  $\mathcal{A}_I(2)$  through  $\mathcal{A}_I(6)$  yield the same period-II welfare as does  $\mathcal{A}_I(7)$ . Furthermore, it is straightforward to check, in view of *Results 3-5*, that, in terms of period-I welfare, the high-demand countries' most preferred choice is to form an FTA among themselves, whereas, residually, the most preferred choice left for the low-demand countries is to form an FTA among themselves. Hence it follows that, in terms of discounted sum of welfare, no single or joint deviation from  $\mathcal{A}_I(7)$  to any of  $\mathcal{A}_I(2)$  through  $\mathcal{A}_I(6)$  is a profitable proposition for any deviating country or a group of deviating countries. This is true irrespective of the magnitude of the discount factor.

Thus either  $\mathcal{A}_I(7)$  or  $\mathcal{A}_I(8)$ , which is multilateralism, can be the PCPNE. Let us now compare these. *Result 4* readily implies that the period-I welfare is higher in the former (as multilateralism entails unilateralism in period I). Considering welfare in period II, if  $\rho \leq \bar{\rho}$ , there is no difference between  $\mathcal{A}_I(7)$  and  $\mathcal{A}_I(8)$  (global free trade will prevail in period II). Thus, in terms of discounted sum of welfare  $\mathcal{A}_I(7)$  strictly dominates  $\mathcal{A}_I(8)$  and hence is the PCPNE.

Suppose  $\rho > \bar{\rho}$ . In this case, for the high-demand countries, one-period welfare under polarization exceeds that under global free trade (*Result 2*). Thus their period-II welfare will be higher under polarization than under global free trade. This, together with that period-I welfare is higher for any country, implies that the high-demand countries' discounted sum of welfare is greater in  $\mathcal{A}_I(7)$  than in  $\mathcal{A}_I(8)$ . Hence these countries surely block multilateralism. Low-demand countries may or may not want to block it as their period-I welfare is higher under polarization but their period-II welfare is less. However this is immaterial since multilateralism is blocked by the high-demand countries anyway.

Thus  $\mathcal{A}_I(7)$  is the unique PCPNE. ■

In summary, note that the relative market access effect and the relative competition effect, which dictate *Results 1-5*, explain why  $\mathcal{A}_I(7)$  is the solution.

### 3.3 Other Related Results

There are other related issues, which our model throws some light upon – such as regionalism versus multilateralism, who may oppose to regionalism or multilateralism and what happens to equilibrium tariffs when PTAs are formed.

#### Regionalism versus Multilateralism

As an immediate corollary of Proposition 3,

**Proposition 4** *In our model economy, multilateralism is blocked under any parametric configuration.*

### Who Blocks Multilateralism?

Let us interpret  $\mathcal{A}_I(7)$  as regionalism. In course of the proof of Proposition 3 above, we have seen that high-demand countries are always better off under regionalism than under multilateralism. Hence they block global free trade under any parametric configuration. To a low-demand country period-I welfare is higher under regionalism whereas period-II welfare may be less. It is then possible that low-demand countries want to block multilateralism too.

If  $\rho \leq \bar{\rho}$ , for any country, regionalism yields the same period-II welfare as does multilateralism. Hence the low-demand countries would block multilateralism too. When  $\rho > \bar{\rho}$ , compared to multilateralism, a low-demand country's period-II welfare is less. The discounted sum of welfare of country  $C$ , for example, under the alternative systems are given by:

$$\begin{aligned} \text{Multilateralism: } W_M^C &= W_U^C + \delta W_F^C = \frac{101a_L^2 + 4a_H^2}{242} + \delta \left( \frac{10a_L^2 + 2a_H^2}{25} \right), \\ \text{Regionalism: } W_R^C &= (1 + \delta)W_{fta}^C = \frac{14a_L^2 + a_H^2}{32}(1 + \delta). \end{aligned}$$

Comparing the two expressions,  $W_R^C > W_M^C$  iff

$$\begin{aligned} \frac{14a_L^2 + a_H^2}{32}(1 + \delta) &> \frac{101a_L^2 + 4a_H^2}{242} + \delta \left( \frac{10a_L^2 + 2a_H^2}{25} \right), \\ \iff 1950 + 3630\delta + (1425 - 4719\delta)\rho^2 &> 0. \end{aligned} \tag{9}$$

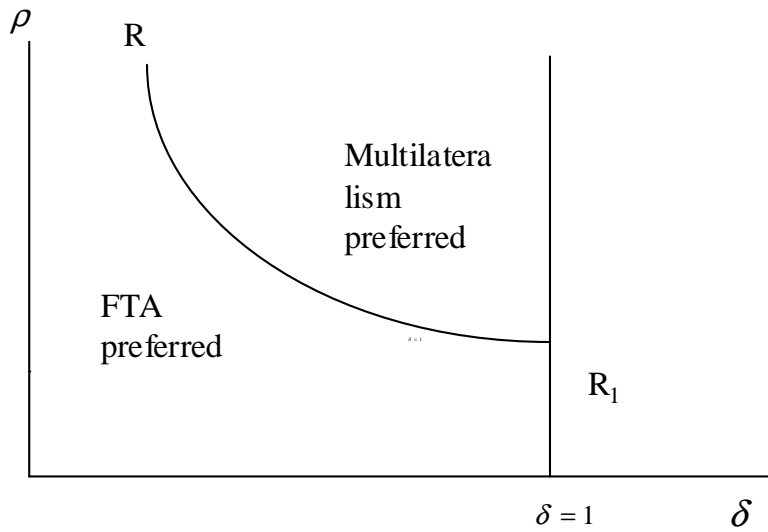


Figure 2: Low-demand Countries' Preference to Support Regionalism or Multilateralism



The above inequality can be characterized in the  $(\rho, \delta)$  space by the equation

$$1950 + 3630\delta + (1425 - 4719\delta)\rho^2 = 0, \quad (9')$$

which is graphed as the  $RR_1$  curve in Figure 2. The inequality (9) is satisfied below the  $RR_1$  schedule. It then follows that

**Proposition 5** *The high-demand countries block multilateralism under any parametric configuration. The low-demand countries want to block multilateralism if and only if  $(\rho, \delta)$  lie below the  $RR_1$  schedule in Figure 2.*

### 3.3.1 Comparison of Tariff Rates under Polarization and Unilateralism

In Krugman (1991), the tariffs rise with the formation of trading blocs. However, under GATT Article XXIV, PTA formation is permissible only when the average tariffs against non-members are not raised above the pre-existing rates. Interestingly, in our model, comparing the tariff rates under unilateralism spelled in (6) and (7) with those under polarization, as given in (A3) in Appendix 1, we find that

**Proposition 6** *The tariff rates in the polarized world economy are lower than those under unilateralism.*

This is consistent with the aforementioned GATT article. The reason behind falling tariffs is the ‘complementarity effect’ of tariff reduction – when some countries form an FTA, imports from non-member countries fall, which tend to lower a member country’s market power vis-a-vis the nonmember countries. As a result, member countries lower their tariffs against nonmembers. The complementarity effect is not new: indeed it is noted in Freund (2000). However, Freund’s analysis assumed that all trading countries are similar. Here it is shown that this effect holds even when countries are heterogeneous.

## 4 Single Period: All forms of negotiations have the same transaction cost

Our assumption that multilateral negotiations take longer to negotiate and implement than FTAs was meant to capture the notion that the former entail higher transaction costs than do the latter. Suppose instead that all types of negotiations and implementation take one period. (This assumption suppresses the dynamic time path issue). In this case, whatever negotiations that take place between countries is implemented at the beginning of period I, and continues in period II. Hence, with no loss of generality,

we can consider the equilibrium FTA formation in a single period framework. The equilibrium notion here is then CPNE.

There are two essential differences from the earlier model. One is that multilateralism can be a CPNE. Second, global free trade and multilateralism mean the same (as global free trade via regionalism is ruled out by assumption).

In terms of various possibilities, in Table 1, column 2 disappears and in column 1 the last entry is global free trade meaning the bloc,  $\{A-B-C-D\}$ . What is the CPNE? Our discussion of period-II CPNE in the previous model leads to

**Proposition 7** *When all types of negotiations are completed instantaneously, the CPNE is either polarization or global free trade according as  $\rho \geq \bar{\rho}$ . The global free trade may be blocked by North countries only and they do block it if  $\rho > \bar{\rho}$ .*

Again, the relative market access effect and the relative competition effect explain this result. Figure 3 illustrates it.

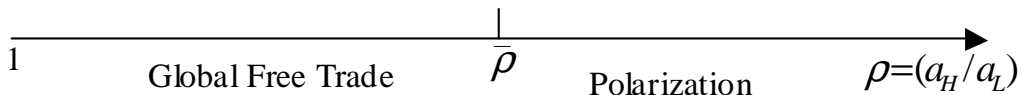


Figure 3: Ranges of  $\rho$  where Polarization or Global Free Trade Occurs

We now examine two important extensions of this basic set up.

#### 4.1 Difference in Market Structure

An assumption maintained till now is that all countries, irrespective of their market sizes, have the same number of firms (normalized to one for simplicity). Suppose that the number of firms is greater in some countries than in some others. This adds a source of dissimilarity (or comparative advantage). Yet, remarkably, the polarization outcome continues to emerge. Further, compared to when there is no difference in market structure, the scope of global free trade is even less.

Of course, the most general way to incorporate heterogeneous number of firms will be to allow free entry and exit such that the number of firms in each country is determined by zero profit in the *entire* global market. However, such a model does not seem to be analytically tractable, because profits across markets are no longer independent of one another. The next best alternative is to assume that different countries have different number of firms, arbitrarily given, as in Krishna (1998). This is also analytically daunting.

In what follows, we consider a more specific situation where one set of countries have one firm each and the other set of countries have two firms each. Appendix 2 works out this model. As shown there,

if an FTA structure is the CPNE, it is in the form of polarization. The economic factors underlying this result do however not seem to critically depend on our particular assumption on the number of firms – which suggests that this is likely to hold for arbitrarily given number of firms in each set of countries. This is significant, because the difference in market structure is a source of comparative advantage on its own, and, yet there is no profitable single or joint deviation from PTAs among the same-market-size countries.

There are two alternative cases in our worked-out model: (a) two firms in each high-demand country and one firm in each low-demand country and (b) vice versa, i.e., one firm in each high-demand country and two firms in each low-demand country.

Consider first the case (a). As before, polarization occurs when  $\rho$  is large enough. The reason is the same as before: the high-demand countries block global free trade and match with each other. Interestingly however, polarization occurs when the market size difference is sufficiently small too. In this case it is low-demand countries who block global free trade. Given that  $\rho$  is sufficiently close to one, by a move from polarization to global free trade, these countries get access to only a slightly bigger market as a whole, while they give freer access to a relatively large number of firms in their own markets. The relative market access effect and the relative competition effect imply that their profits are less under global free trade than under polarization. This is why they block global free trade.<sup>15</sup> Global free trade emerges as equilibrium only over an intermediate range of  $\rho$ . More precisely,

**Proposition 8** *Given that each high-demand country has two firms and each low-demand country has one, global free trade or polarization is the CPNE according as  $\rho \in (\rho_1, \rho_2)$  or  $\rho \notin (\rho_1, \rho_2)$ , where  $\rho_1 \equiv 1.0873 < \bar{\rho}$  and  $\rho_2 \equiv 1.338 > \bar{\rho}$ . In the range,  $\rho > \rho_2$  global free trade is blocked by the North countries, whereas in the range  $\rho < \rho_1$  it is blocked by the South countries.*

This is illustrated in Figure 4.

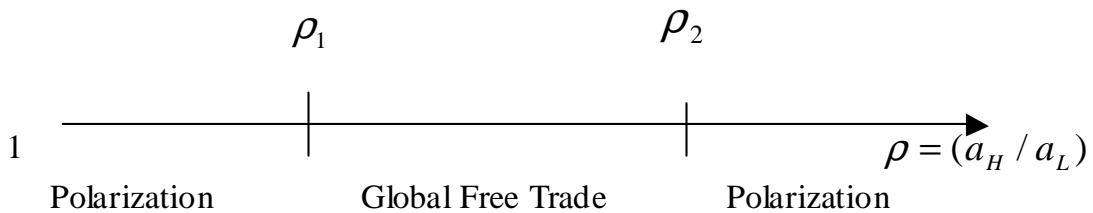


Figure 4: Ranges of  $\rho$  where polarization and global free trade occur

Let us now characterize case (b). Even a stronger result holds:

<sup>15</sup>Although their consumer surplus increases, it is dominated by the fall in profits.

**Proposition 9** *If each high-demand country has one firm and each low-demand country has two, polarization is the CPNE for all values  $\rho$ . It is the North countries who block global free trade for all parametric configurations in this case.*

Therefore, global free trade is not an equilibrium outcome under any (legitimate) parametric configuration. Intuitively, when country sizes are sufficiently different, the North countries block global free trade because of both the relative market access effect and the relative competition effect. This is because they have sufficiently bigger market size as well as lesser number of firms. Here both these effects are strong and reinforce each other.

When country sizes are similar, as seen in case (a), the countries with less number of firms – in this case the high-demand ones – prefer polarization, and block global free trade. This, as mentioned before, is the relative competition effect.<sup>16</sup> In other words, in case (b), the high-demand countries block global free trade in preference to polarization under *all* parametric configurations.

The low-demand countries *always* prefer global free trade to polarization in this case, since both the relative market access effect and the relative competition imply higher profits and welfare under the former. However, since the high-demand countries block global free trade, they match with each other residually.

## 4.2 More Countries

Finally, we relax the assumption that there are only four countries. Suppose that there are more than four countries, and further the number of low-demand countries is greater than that of the high-demand countries.<sup>17</sup> Formalizing this in terms of arbitrary number of countries does not again seem tractable. But, as an example, Appendix 3 works out a *five*-country model with two high-demand and three low-demand countries. In this case, there are 16 possible coalition structures, including global free trade, as opposed to 9 in the 4-country model. Reverting back to the assumption that there is one firm in each country, the comparison between these options yield that the CPNE is still either polarization if  $\rho$  is above a critical value, or global free trade otherwise. More precisely,

**Proposition 10** *If there are two-high-demand and three-low-demand countries, polarization or global free trade is the CPNE, as  $\rho \gtrless \rho_3 \equiv 1.068$ . It is only the high-demand countries who might block global free trade and they do when  $\rho > \rho_3$ .*

When the combined market size of the North countries is bigger than that of the South countries, i.e.  $2a_H > 3a_L$ , the former prefer polarization because of the relative market access effect. Since the three South countries together have more firms than the two North countries, the relative competition

---

<sup>16</sup>The relative market access effect is negligible here, since each country has almost similar market size

<sup>17</sup>There are many more developing countries than developed countries in the world economy.

effect reinforces the preference of the North countries to polarize, rather than move to global free trade. Otherwise, when  $3a_L > 2a_H$ , there are relative market access gains from moving to global free trade for the North countries. Yet, the relative competition effect ensures that Northern countries may still prefer polarization, unless the size difference between a North and a South country is sufficiently small, namely  $\rho < \rho_3$ . When  $\rho < \rho_3$ , the relative market access gains dominate the relative competition losses so that the North countries also prefer global free trade (the South countries *always* prefer global free trade to polarization).<sup>18</sup>

These two effects also ensure why, from polarization, the North countries are worse off under any three-country or four-country FTA, and hence block such a deviation. Therefore, the CPNE is either global free trade or polarization, as in the four country case. The low demand countries always prefer global free trade to polarization, but when the high-demand countries block global free trade their best option is to match residually.

The upshot is that the relative market access effect and the relative competition effect govern the outcomes once again. The general nature of these effects suggests that our result is likely to hold when there are arbitrary numbers of high-demand and low-demand countries.

## 5 Concluding Remarks

This paper formulates a model of an asymmetric global economy, in which there are at least four countries. This enables us to address the hitherto unexplored issue of whether trading blocs will consist of similar or dissimilar countries. Also it examines the dynamic time path issue – how regional blocs may or may not lead to global free trade. The countries are primarily differentiated on the basis of market size. The difference in market structure is also considered.

The main result of the paper is that *if FTA is the equilibrium outcome, it is in the form of polarization*. As shown, it does not just arise from the usual gains from trade among similar countries. Because it holds when countries differ in size as well as market structure. Two effects, namely, a relative market access effect and a relative competition effect, are shown to be the key behind the result.

Given that the developed and developing countries have some basic dissimilarities, the model may be interpreted as an attempt to examine whether the developed (developing) countries form trading blocs with other developed (developing) countries, or whether blocs between developed and developing countries emerge. As noted in the Introduction, majority of trading blocs to date are between similar countries. Trading blocs with North-South mixing like NAFTA are, as of now, more of an exception than the rule. Thus the polarization result of this paper is consistent with this overall pattern.

---

<sup>18</sup>Note that  $\rho_3 < \bar{\rho}$ , i.e. the parametric configurations where polarization occurs here is a superset of that in our earlier four-country model.

The model also predicts that equilibrium tariffs fall as PTAs are formed. Furthermore, in an asymmetric world economy some countries may block global free trade. Within the purview of the model we have identified who may block it and under what parametric configurations.

Although till now a vast majority of trading blocs are among similar countries – which is the focus of this paper – North-South blocs are a growing phenomenon. The obvious next task will be to identify conditions under which such blocs may arise. A general equilibrium model, perhaps similar to that of Levy (1997), accommodating comparative advantage by different countries in different sectors together with factors that lead to gains from trade among similar countries will have the scope to predict such blocs.

North-South trading blocs have also been predicted by Ethier (1998) as well as Arias and Spiegel (1998). However, in both these papers the rationale for such blocs stems from FDI from North to South. A natural extension will be allow for FDI in the framework used in this paper, and identify conditions under which North-South trading blocs may arise.

## REFERENCES

- Arias, E. F. and M. M. Spiegel, 1998. North-South Customs Unions. *Journal of International Economics*, 46, 229-251.
- Bagwell, K. and R. Staiger. 1997. Multilateral Tariff Cooperation during the Formation of a Customs Union. *Journal of International Economics*, 42. 91-123.
- Baldwin, R. 1995. A Domino Theory of Regionalism. *Baldwin, R., P. Haaparanta and J. Kiander (eds.), Expanding Membership of the European Union (Cambridge University Press)*.
- Bhagwati, J. 1993. Regionalism and Multilateralism: An overview. *de Melo, J. and A. Panagariya (eds.) New Dimensions in Regional Integration (World Bank and Cambridge University Press)*. 22-51.
- Bhagwati, J. and A. Panagariya, 1996. Preferential Trading Areas and Multilateralism: Strangers, Friends or Foes?. *Bhagwati, J. and A. Panagariya (eds.), The Economics of Preferential Trade Agreements (AEI Press, Washington DC)*. 83-100.
- Bond, E., C. Syropoulos, and L.A. Winters, 2001. Deepening of Regional Integration and Multilateral Trade Agreements. *Journal of International Economics* 53. 335-362.
- Brander, J. and P. Krugman, 1983. A 'Reciprocal Dumping' Model of International Trade. *Journal of International Economics* 15. 313-323.
- Burbridge, J., J. DePater, G. Myers and A. Sengupta, 1997. A Coalition-Formation Approach to Equilibrium Federations and Trading Blocs. *American Economic Review*, (December). 940-956.
- Clark, D. P. and D. L. Stanley, 1999. Determinants of Intra-Industry Trade Between Developing Countries and the United States. *Journal of Economic Development*, 24 (2), 79-94.
- Ethier, W. 1998. Regionalism in a Multilateral World. *Journal of Political Economy*, 106 (6). 1214-1245. 1-22.
- Fernandez-Arias Eduardo and Mark M. Spiegel, 1998. North South Customs Unions and International Capital mobility. *Journal of International Economics* (46)2 229-251.
- Freund, C. 2000. Multilateralism and the endogenous formation of preferential trade agreements. *Journal of International Economics* (52)2 359-376.
- Freund, C. 2000. Different Paths to Free Trade: The Gains From Regionalism. *Quarterly Journal of Economics* 15 (4). 1317-1341.
- Fudenberg, D. and D. Levine, 1983. Subgame-Perfect Equilibria of Finite and Infinite Horizon Games. *Journal of Economic Theory*, 31. 251-268.
- Kandogan, Yener, 2003. "Intra-industry Trade of Transition Economies: Trends and Determinants", Working Paper No. 2003-01, School of Management, University of Michigan-Flint.
- Krishna, P. 1998. Regionalism and Multilateralism: A Political Economy Approach. *Quarterly Journal of Economics*, 113 (1). 227-251.
- Krugman, P. 1991. Is Bilateralism Bad. *Helpman, E. and A. Razin (eds.), International trade and Trade Policy. M.I.T. Press, Cambridge MA*
- Lipsey, R. G. 1957. The Theory of Customs Unions: Trade Diversion and Welfare. *Economica*, 24. 40-46.
- Levy, P. 1997. A Political Economy Analysis of Free-Trade Agreements. *American Economic Review*, 87. 506-519.

- Maskin, E. and J. Tirole. 2001. Markov Perfect Equilibrium, I: Observable Actions. *Journal of Economic Theory* 100, 191-219.
- Meade, J.E. 1955. *The Theory of Customs Unions*, Amsterdam: North Holland. item Nilsson, L. 1997. *Essays on North South Trade*, Doctoral Dissertation, Lund University, Sweden.
- Panagariya, A. 2000. Preferential Trade Liberalization: The Traditional Theory and New Developments. *Journal of Economic Literature*, XXXVIII. 287-331.
- Peleg, B. 1992. On Perfectly Coalition Proof Nash Equilibria. *Majumdar, M. (ed), Equilibrium and Dynamics: Essays in Honor of David Gale*. New York: St. Martins Press.
- Ray, D. 1998. *Development Economics*. Oxford University Press.
- World Bank. 2000. *Trade Blocs. Policy Research Report*, Washington.
- Viner, J. 1950. The Customs Unions Issue. *Carnegie Endowment for International Peace*.
- Yi, S-S. 1996. Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good. *Journal of International Economics* 41. 153-177.
- Yi, S-S. 2000. Free Trade Areas and Welfare: An Equilibrium Analysis. *Review of International Economics*, May. 153-177.



## Appendix 1: Proofs of Propositions 1 and 6 and Results 2-3

We first compute the free trade welfare levels. By setting the tariffs to zero in eq. (5) in the text,

$$W_F^A = W_F^B = \frac{10a_H^2 + 2a_L^2}{25}; \quad W_F^C = W_F^D = \frac{10a_L^2 + 2a_H^2}{25}. \quad (\text{A1})$$

We now obtain the welfare levels of the high and low demand countries under polarization. In country  $A$  ( $B$ ) the firm from  $B$  ( $A$ ) enjoys tariff-free access. This yields the following output solutions in stage 2 (in country  $A$ ):

$$z_A^A = z_B^A = \frac{a_H + 2t^A}{5}; \quad z_C^A = z_D^A = \frac{a_H - 3t^A}{5}.$$

The profit expressions are simply the square of the quantity expressions. Using the output and profit expressions above, the welfare of country  $A$  under polarization is given by

$$\begin{aligned} W_P^A = & a_H \left( \frac{4a_H - 2t^A}{5} \right) - \frac{1}{2} \left( \frac{4a_H - 2t^A}{5} \right)^2 - 2 \left( \frac{a_H - 3t^A}{5} \right)^2 \\ & - \left( \frac{a_H + 2t^A}{5} \right)^2 + \left( \frac{a_H + 2t^B}{5} \right)^2 + \left( \frac{a_L - 3t^C}{5} \right)^2 + \left( \frac{a_L - 3t^D}{5} \right)^2. \end{aligned} \quad (\text{A2})$$

Maximizing this with respect to  $t^A$ , and, by virtue of symmetry, we obtain  $t_P^A$  and  $t_P^B$ , where  $P$  stands for polarization.  $t_P^C$  and  $t_P^D$  are obtained likewise. We have

$$t_P^A = t_P^B = \frac{a_H}{8}; \quad t_P^C = t_P^D = \frac{a_L}{8}. \quad (\text{A3})$$

Comparing these expressions proves Proposition 6.

By substituting the tariff expressions from (A3) in the welfare expression of a high-demand country in (A2), the one-period welfare of such a country under polarization is obtained, while that of a low demand country is derived analogously. We have,

$$W_P^A = W_P^B = \frac{14a_H^2 + a_L^2}{32} > W_U^A = W_U^B, \quad W_P^C = W_P^D = \frac{14a_L^2 + a_H^2}{32} > W_U^C = W_U^D. \quad (\text{A4})$$

By comparing these welfare levels with that under global free trade, given in (A1),

$$\begin{aligned} W_F^C &> W_P^C \text{ for any } \rho > 1, \\ W_F^A &> W_P^A \text{ iff } \rho < 1.14 \equiv \bar{\rho}. \end{aligned} \quad (\text{A5})$$

Thus, the low demand countries are *always* better off under global free trade than under polarization, and, the high demand countries better or worse off as  $\rho$ , the market size ratio is less than or greater than a critical value. This proves *Result 2*.

Now we use these welfare expressions to compare with the welfare levels of a high (low) demand country under other possible coalitions. The latter can be derived in a similar fashion. In what follows, we explicitly state only the welfare expressions under other possible coalitions, while their derivations are skipped. Consider the set of all coalition partitions possible in this 4-country model. They are the following.

- (i)  $\{A, B, C, D\}$ : all countries remain unilateral,
- (ii)  $\{A-B, C-D\}$ : polarization,
- (iii)  $\{A-i, B-j\}$ , where  $i$  and  $j$  are low-demand countries: 2 North-South PTAs,
- (iv)  $\{A-B, C, D\}$ : high-demand countries match, but low-demand ones stay unilateral,
- (v)  $\{A, B, C-D\}$ : low-demand countries match, but high-demand ones stay unilateral,
- (vi)  $\{i-j, k, l\}$ ,  $i$  and  $k$  ( $j$  and  $l$ ) are high- (low-) demand countries: only one high-demand and one low-demand country match,

- (vii)  $\{A-B-i, j\}$ , where  $i$  and  $j$  are low-demand countries,
- (viii)  $\{C-D-k, l\}$ , where  $k$  and  $l$  are high-demand countries,
- (ix)  $\{A-B-C-D\}$ : global free trade.

Note that this exhausts the set of all possible coalition structures that may occur in our model economy. However, all of these coalitions can occur corresponding to the action-sets  $\mathcal{A}_I(2)$ - $\mathcal{A}_I(7)$  in period I. Corresponding to the remaining action set, namely  $\mathcal{A}_I(1)$ , i.e. when all countries remain unilateral in period I, global free trade cannot be negotiated in period II, by assumption.

We first compute the CPNE corresponding to  $\mathcal{A}_I(7)$ , that is, a period I action-set from all the nine possible coalitions depicted above can occur. Here the CPNE is polarization if  $\rho > \bar{\rho}$ , and global free trade otherwise. Our task is to show, that there are no possible ‘profitable’ deviation by a single country or a group of countries from it.

First, comparing the welfare levels under polarization given in (A5) with that under unilateralism as given in equations (8) in the main text, we find that the welfare under polarization is higher than the welfare levels under unilateralism for *any* country. Hence unilateralism is not a profitable deviation for any country: it is ‘strongly’ dominated.

Similarly, if a single country deviates from polarization by moving to unilateralism, the welfare of each country goes down. This rules out deviations as given in action sets  $\mathcal{A}_I(2)$ - $\mathcal{A}_I(4)$  in Table I, namely  $(A-B), (C), (D), (A), (B), (C-D)$  and  $(A-C), (B), (D)$ .

Let us now consider two-way FTA jumping from polarization, i.e. formation of North-South FTAs  $(A-C), (B-D)$ . It is found that both a high-demand country, say  $A$  and a low-demand country, say  $B$  is better off forming an FTA with  $B$  rather than a low-demand country. Formally,

$$\begin{aligned} W_P^A &= \frac{14a_H^2 + a_L^2}{32} > W_{(A-C), (B-D)}^A = \frac{25a_H^2 + 5a_L^2}{64} \\ W_P^C &= \frac{14a_L^2 + a_H^2}{32} < W_{(A-C), (B-D)}^C = \frac{25a_L^2 + 5a_H^2}{64}. \end{aligned} \quad (\text{A6})$$

Equation (A6), along with the fact that polarization higher welfare than unilateralism proves *Result 3*.

Further, equation (A6) implies that North-South mixing through ‘two-way FTA jumping’ of this sort is not a profitable deviation for the high-demand countries from polarization, and hence they will oppose such a deviation.

Consider now one-way FTA jumping, which is of two types: (a) a low-demand country joins the already formed FTA between the two high-demand countries, i.e. FTA partitions like  $\{(A-B-C), (D)\}$ , and (b) a high-demand country joins the already formed FTA between the two low-demand countries, i.e. partitions like  $\{(A-C-D), (B)\}$ .

It is found that

$$W_P^A > W_{(A-B-C), (D)}^A = \frac{68183.5a_H^2 + 9113a_L^2}{(37 \cdot 11)^2}, \quad (\text{A7})$$

that is the high-demand countries is better off in polarization than by including any single South country. Also,

$$W_P^A > W_{(A-C-D), (B)}^A = \frac{61880.5a_H^2 + 15488a_L^2}{(37 \cdot 11)^2}, \quad (\text{A8})$$

that is, it is better for a high demand country to stay polarized than to form a FTA with two low-demand countries. In other words, one-way FTA jumpings will make the partner high-demand country worse off, and hence it will block such a proposed deviation.

Hence the welfare of the high demand countries are *always* higher in polarization, that is, in  $\{(A-B), (C-D)\}$  compared to any other FTA structure, except for global free trade. As such they will *always* block any proposed deviation from polarization. The high demand countries obtain higher welfare under global free trade than under polarization if  $\rho \leq \bar{\rho} \equiv 1.14$ , otherwise they are still better off under polarization. Hence the CPNE corresponding to  $\mathcal{A}_I(6)$  will be polarization or global free trade according as  $\rho \geq \bar{\rho}$ .

Since the possible coalitions from  $\mathcal{A}_I(7)$  is the set of all feasible coalitions that may be formed in period II, the CPNE corresponding to it will be the CPNE in  $\mathcal{A}_I(2) - \mathcal{A}_I(6)$  as well, when polarization

and global free trade are both attainable.

However, global free trade is not attainable under  $\mathcal{A}_I(1)$ . In this case,  $\{(A-B), (C-D)\}$  i.e. polarization is feasible, and since it has been shown that it ‘weakly’ dominates the other possible FTA structures except global free trade, (in the sense that the high-demand are always better off under it) it will be the CPNE.

Proposition 1 is thus proved.

## Appendix 2: Difference in Market Structure

### High demand countries have more firms

Here the countries  $A$  and  $B$  have two firms each, while the countries  $C$  and  $D$  have one each. The number of possible coalitions here, as in our basic model, is 9, among which we need to identify the CPNE.

First, we consider unilateralism: the status quo. Proceeding exactly as in our basic model, the equilibrium tariffs for a high-demand and a low-demand country are respectively  $t_H = \frac{5a_H}{22}$  and  $t_L = \frac{3a_L}{13}$ . Using these, the respective welfare expressions are:

$$W_U^A = W_U^B = \frac{222a_H^2}{22^2} + \frac{4a_L^2}{169}; \quad W_U^C = W_U^D = \frac{145a_L^2}{338} + \frac{a_H^2}{242}. \quad (\text{A9})$$

We now do the same assuming polarization, and, obtain

$$W_P^A = W_P^B = \frac{17a_H^2}{36} + \frac{a_L^2}{25} > W_U^A = W_U^B, \quad W_P^C = W_P^D = \frac{11a_L^2}{25} + \frac{a_H^2}{72} > W_U^C = W_U^D. \quad (\text{A10})$$

Thus polarization dominates unilateralism.

We now claim that the CPNE coalition structure here is global free trade if ( $1.087 \equiv \rho_1 \leq \frac{a_H}{a_L} \leq 1.338 \equiv \rho_2$ ), or polarization otherwise. To prove the claim, it is shown that there exists no unilateral or joint profitable deviation by any country or groups of countries from the CPNE.

First, note that similar to unilateralism, in case of deviations like  $\{(A-B)(C), (D)\}$ , or  $\{(A), (B), (C-D)\}$  where one set of countries stays polarized, while the other set deviates by moving to their unilateral position, the deviating countries earn lower welfare than polarization. Hence this also is not a profitable deviation.

Consider a two-way deviation by the countries to two ‘North-South’ FTAs, i.e. coalitions like  $\{(A-C)(B-D)\}$ . It is found that

$$\begin{aligned} W_{\{(A-C)(B-D)\}}^A &= \frac{1471a_H^2}{3362} + \frac{178a_L^2}{2209} > W_P^A \text{ if } \rho < 1.082 \\ W_{\{(A-C)(B-D)\}}^C &= \frac{1789a_L^2}{4418} + \frac{73a_H^2}{1681} > W_P^C \text{ if } \rho > 6.12. \end{aligned} \quad (\text{A11})$$

Thus a high-demand country prefers to form an FTA with a low demand country only if the market size is sufficiently similar. A low-demand country, however, prefers to form an FTA with a high-demand country only if the market size is sufficiently different. However, there is no such parametric configuration where a high-demand country wants to form an FTA with a low-demand country, and it is *reciprocated*. As a result, such a coordinated joint deviation from polarization will never occur.

Next, we consider one-way deviations from polarization. First consider FTA partitions like  $\{(A-C-D)(B)\}$ , i.e. the option of a high-demand country to form an FTA with both the low-demand countries rather than the other high demand country. The corresponding welfare expressions for the high and low demand countries are respectively as follows:

$$\begin{aligned} W_{\{(A-C-D)(B)\}}^A &= \frac{611a_H^2}{1425} + \frac{25a_L^2}{256} < W_P^A \text{ if } \rho > 1.07, \\ W_{\{(A-C-D)(B)\}}^C &= \frac{425a_L^2}{1024} + \frac{65a_H^2}{2232} < W_P^C \text{ if } \rho < 1.25. \end{aligned} \quad (\text{A12})$$

This implies that in the parametric configurations where a high-demand country prefers this option to

polarization, the low demand countries prefer to remain polarized rather than include a high-demand country. Hence, under no parametric configurations will this coalition block polarization.

Now let us compare the option of polarization with that of the other one-way deviation, i.e. coalition structure  $\{(A-B-C), (D)\}$ , i.e. if two high-demand and one low-demand country form a FTA, with the other low-demand country remaining isolated. The corresponding welfare expressions are:

$$\begin{aligned} W_{\{(A-B-C), (D)\}}^A &= \frac{5711a_H^2}{12482} + \frac{6866a_L^2}{123201} < W_P^A \text{ if } \rho > 1.07 \\ W_{\{(A-B-C), (D)\}}^C &= \frac{97281a_L^2}{246402} + \frac{288a_H^2}{6241} < W_P^C \text{ if } \rho < 1.19. \end{aligned} \quad (\text{A13})$$

Therefore, in this case also, in those parametric configurations where the two high-demand countries will prefer to form a FTA with a low-demand country the low-demand country will not find it optimal to join: it will prefer polarization. Therefore, under no circumstances will this coalition block polarization.

Finally we compute the welfare of the high demand and the low demand countries in global free trade. The respective welfare levels are

$$W_G^A = \frac{22a_H^2 + 4a_L^2}{49}; \quad W_G^C = \frac{20a_L^2 + 2a_H^2}{49}.$$

It is found that  $W_G^C > W_P^C$  if  $\rho > 1.087$ , while  $W_P^A > W_G^A$  if  $\rho > 1.338$ .

Therefore, the CPNE coalition structure here is global free trade if  $(1.087 \equiv \rho_1 \leq \frac{a_H}{a_L} \leq 1.338 \equiv \rho_2)$ , or polarization otherwise. In other words, there will be global free trade if the market size ratio lies in an intermediate range. If it is too high, the high-demand countries will prefer polarization and block global free trade, while if it is too low, the low-demand countries will do the same.

### Low-demand countries have more firms

In this case, the welfare under unilateralism of a high demand and a low demand country is the following:

$$W_U^A = W_U^B = \frac{145a_H^2}{338} + \frac{a_L^2}{242}; \quad W_U^C = W_U^D = \frac{222a_L^2}{222} + \frac{4a_H^2}{169}. \quad (\text{A14})$$

The welfare expression under polarization is as follows:

$$W_P^A = \frac{11a_H^2}{25} + \frac{a_L^2}{72} > W_U^A; \quad W_P^C = \frac{17a_L^2}{36} + \frac{a_H^2}{25} > W_U^C. \quad (\text{A15})$$

Thus welfare under polarization is higher than that under unilateralism for *both* countries. Being a four-country model, here also the possible FTA partitions is 9, which is spelled out in Appendix 1. Among them, the CPNE here is polarization for *all* parametric configurations. To prove it, as before, consider the possible deviations from polarization, which we need to show is unprofitable.

It has been already found that the welfare of em any country is higher under polarization than under unilateralism, and is hence not a profitable deviation. Similarly, in case of deviations like  $\{(A-B), (C), (D)\}$ , or  $\{(A), (B), (C-D)\}$  where one set of countries stays polarized, while the other set deviates by moving to their unilateral position, the deviating countries earn lower welfare than polarization. Hence this is not a profitable deviation as well.

Consider now two-way deviations from polarization i.e. ‘North-South’ coalition partitions like  $\{(A-C)(B-D)\}$ . It is found that

$$W_{\{(A-C)(B-D)\}}^A = \frac{1789a_H^2}{4418} + \frac{73a_L^2}{1681} < W_P^A; \quad W_{\{(A-C)(B-D)\}}^C = \frac{1471a_L^2}{3362} + \frac{178a_H^2}{2209} > W_P^C. \quad (\text{A16})$$

In other words, both the high demand and the low demand country prefers the other high-demand country as its trading partner, and hence though the low-demand countries will prefer such a deviation, the high demand countries will oppose and block such a proposed deviation from polarization.

Let us now consider one-way deviations. First consider the coalition structure  $\{(A-C-D), (B)\}$ , i.e.

when a high-demand forms an FTA with both the low demand countries, with the other high demand country remains isolated. It is obtained that

$$W_{\{(A-C-D)(B)\}}^A = \frac{97281a_H^2}{246402} + \frac{288a_L^2}{6241} < W_P^C \text{ always.} \quad (\text{A17})$$

Thus a high-demand country is better off forming an FTA with the other high-demand country rather than both the low-demand countries, and will block the above possible deviation.

We also obtain that a high-demand country is better off in polarization rather than allowing a single low-demand country to join their FTA. In other words, it will block one-way deviations of the form of  $\{(A-B-C), (D)\}$ . Thus

$$W_{\{(A-C), (D)\}}^A = \frac{425a_H^2}{1024} + \frac{65a_L^2}{2232} < W_P^A. \quad (\text{A18})$$

Finally the welfare under global free trade of the high-demand and the low-demand countries are stated as:

$$W_G^A = \frac{20a_H^2 + 2a_L^2}{49} < W_P^A \text{ always; } W_G^C = \frac{22a_L^2 + 4a_H^2}{49} > W_P^C \text{ always.} \quad (\text{A19})$$

Therefore, while the low-demand countries prefer global free trade to polarization, the high-demand countries do not and hence they will block any proposal by the low-demand countries toward moving to global free trade from polarization. Polarization is the only CPNE coalition structure for *all* parametric configurations.

### Appendix 3: Five Countries

A single period model with two high-demand ( $A$  and  $B$ ) and three low-demand countries ( $C$ ,  $D$  and  $E$ ) is considered. First, as before, we consider the system under unilateralism. Proceeding exactly as in the 4 country model, it is found that the equilibrium tariffs for a high demand and a low demand country under unilateralism is  $t_H = \frac{a_H}{4}$  and  $t_L = \frac{a_L}{4}$ . Using this, we obtain that the respective welfare expressions as follows:

$$W_U^A = W_U^B = \frac{61a_H^2 + 3a_L^2}{144}; \quad W_U^C = W_U^D = W_U^E = \frac{62a_L^2 + 2a_H^2}{144}. \quad (\text{A20})$$

Consider now the case of polarization, that is when the two high-demand countries form an FTA, and the three low-demand countries do the same. The welfare expressions for a high demand and a low-demand country are stated as

$$W_P^A = \frac{62,662a_H^2 + 6075a_L^2}{142884} \quad W_P^C = \frac{14,400a_L^2 + 784a_H^2}{31,752}. \quad (\text{A21})$$

It can be checked that  $W_P^A > W_U^A$ , for *all* parametric configurations. It can also be similarly checked that given that the high demand countries match, it is optimal for the low demand ones to match as well, and also  $W_P^C > W_U^C$ , for *all* parametric configurations, so that polarization as an outcome dominates unilateralism. Therefore, unilateralism can never be an equilibrium outcome of the game.

There are in total 16 possible FTA (coalition) structures that may emerge in equilibrium, as stated below:

- (i)  $\{(A)(B)(C)(D)(E)\}$ , i.e. Unilateralism. (ii)  $\{(A-B)(C-D-E)\}$ , i.e. polarization
- (iii)  $\{(A-B)(C)(D)(E)\}$ , (iv)  $\{(A)(B)(C-D-E)\}$
- (v)  $\{(A-B)(i-j)(k)\}$ , where  $i, j, k$  are three low demand countries, (vi)  $\{(A)(B)(i-j)(k)\}$
- (vii)  $\{(A-B-i)(j)(k)\}$ , (viii)  $\{(A-B-i), (j-k)\}$ ,
- (ix)  $\{(A-i)(B-j)(k)\}$ , (x)  $\{(A-i)(B-j-k)\}$  or  $\{(B-i)(A-j-k)\}$
- (xi)  $\{(A-i)(B)(j)(k)\}$  or  $\{(B-i)(A)(j)(k)\}$ , (xii)  $\{(A-i)(B)(j-k)\}$  or  $\{(B-i)(A)(j)(k)\}$
- (xiii)  $\{(A)(i)(B-j-k)\}$  or  $\{(B)(i)(A-j-k)\}$ , (xiv)  $\{(A-B-i-j)(k)\}$ ,
- (xv)  $\{(A-C-D-E)(B)\}$  or  $\{(B-C-D-E)(A)\}$  (xvi)  $\{(A-B-C-D-E)\}$ , i.e. global free trade.

Which one of these possible coalition structures will emerge as the CPNE of the game? We claim that the CPNE is polarization iff  $\rho \equiv \frac{a_H}{a_L} > 1.068$ , and global free trade otherwise. For this to be CPNE, there should be no country or a group of countries which can deviate and form some other coalition and be better off. To prove this claim, let us consider possible deviations from this coalition.

The proof is as follows. We will show that welfare of the high-demand countries is higher under polarization compared to any other coalition (except the grand coalition, i.e. global free trade), and hence they will block any proposed deviation, that involves them. They prefer to remain matched with each other. The low-demand countries, among themselves, cannot make a possible deviation that makes them better off – their best strategy is match residually.

To start with, we show that for a high-demand country, polarization yields high welfare than forming a 4-country FTA. A high-demand country can form a 4-country FTA with (a) with 3 low-demand countries, or (b) the other high-demand country and any two low-demand countries, i.e. coalition structures (xv)  $\{(A-C-D-E)(B)\}$ , and (xiv)  $\{(A-B-i-j)(k)\}$ . Between the above two, consider first option (a) of a high demand country  $A$ , that is FTA structure  $\{(A-C-D-E)(B)\}$ , which we denote as  $M$ . It is found that

$$W_M^A = \frac{19969a_H^2 + 4800a_L^2}{51984} < W_P^A \text{ for all parametric configurations.}$$

Therefore country  $A$  will prefer polarization rather than forming an FTA with three low-demand countries. Hence  $M$  cannot be a ‘profitable’ deviation’ from polarization for  $A$ , who will block it.

Next consider option (b), which leads to coalition structure  $\{(A-B-C-D)(E)\}$ . We name this coalition structure as  $Q$ . It is found that

$$W_Q^A = \frac{21,202a_H^2 + 3561a_L^2}{51984} < W_P^A \text{ for all parametric configurations.}$$

Therefore a high-demand country is better off under polarization than in a four-country FTA,

Next, consider deviations involving three country FTAs. Consider the partition  $\{(A-B-C)(D-E)\}$ , i.e. there is one FTA with two high-demand and one low-demand country, and another FTA with the two remaining low-demand countries. It is found that

$$W_{\{(A-B-C)(D-E)\}}^A = \frac{184a_H^2}{441} + \frac{242a_L^2}{3969} < W_P^A \text{ for all parametric configurations.} \quad (\text{A22})$$

Thus the high-demand countries will block any proposal by a single low-demand country to form a FTA with them. Proceeding similarly, it can be found that the welfare of a high demand country under polarization is higher than its welfare in *any* other type of 3-country FTAs as well, so that coalition structures such as  $\{(A-C-D)(B-E)\}$  and  $\{(A-B-C)(D)(E)\}$  will also not occur in equilibrium.

Finally we consider the option of global free trade. We find that

$$W_F^A = \frac{29a_H^2 + 6a_L^2}{72}; \quad W_F^C = \frac{31a_H^2 + 4a_L^2}{72}. \quad (\text{A23})$$

Comparing this with the welfare of the countries under polarization in (A11),  $W_F^C > W_P^C$  *always*, but

$$W_P^A > W_F^A \text{ if } \rho \equiv \frac{a_H}{a_L} > 1.068.$$

This implies that polarization or global free trade is the CPNE according as the market size difference between the countries is greater or less than a critical value.