Do South-South Preferential Trade Agreements Undermine the Prospects for Multilateral Free Trade?

Paul Missios* and Halis Murat Yildiz†

July 12, 2013

Abstract

Following the proliferation of preferential trade agreements (PTAs) between South countries, there has been concern that the formation of these agreements undermines efforts towards multilateral trade liberalization. To examine the static and dynamic implications of these agreements relative to other types, we construct a model with two North countries and two South countries, where the North and South are differentiated by production cost. In particular, we look at the differences in welfare across the various forms of bilateral free trade agreements (FTAs) and customs unions (CUs) and their implications for the stability of multilateral free trade. We show that tariff complementarity is strongest under a North-North FTA and weakest under a South-South CU, leading to highest world welfare for North-North PTAs. However, among FTAs, a South-South agreement is the only bilateral FTA that can sustain multilateral free trade under certain cost conditions, but otherwise may make multilateral free trade less likely than all other possible trade regimes. For CUs, free trade is easier to sustain under no agreement than a North-South or South-South CU, but may more likely under a North-North agreement. With multiple independent bilateral PTAs, mixed type agreements (North-South) make multilateral cooperation more likely to be sustainable than polarized agreements (North-North, South-South).

Keywords: trade agreements; free trade; development; tariffs

*Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. E-mail: pmissios@ryerson.ca.

†Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. E-mail: hyildiz@ryerson.ca.

1 Introduction

Over the last six decades, countries have pursued trade liberalization along several fronts: unilaterally, preferentially with a few partner, and multilaterally within the General Agreement on Tariffs and Trade (GATT), now subsumed by the World Trade Organization (WTO). The Article XXIV of the GATT permits WTO member countries to form preferential trade agreements (PTAs) such as free trade agreements (FTAs) and customs unions (CUs) under which members can grant tariff concessions to each other while they do not extend such concessions to other WTO members. As per the WTO's official website, as of Jan. 2012, the WTO had received notification of 511 such arrangements, of which 370 were notified under Article XXIV. About 319 such agreements are already in force and most WTO members belong to multiple PTAs, and recently one can even observe major PTA groups in discussion with each other regarding mutual liberalization. While FTAs constitute an overwhelming majority of PTAs, the existing CUs involve some of the major economies of the world: for example, the Latin American CU MERCOSUR counts Argentina, Brazil, Paraguay, and Uruguay as its members while the EC (27) – a CU that extends across both goods and services – comprises of most major European economies. While the formation of PTAs has accelerated significantly, there has been little progress towards multilateral trade concessions in the last two decades. Further, many of the more recently signed agreements are between lower-income (South) partners, particularly among countries in Asia and Latin America.¹ This proliferation of South-South PTAs has posed an important concern to many in light of the trade diversion that is expected to occur under these agreements.²

Given that the first Article of the GATT, the most favored nation (MFN) clause, is about non-discrimination and is at the heart of the WTO system (making the discriminatory nature of PTAs appear to be in direct conflict with it), the existence of the Article XXIV has not been without controversy.³ Another important exception to non-discrimination is the Enabling Clause that allows developed members to give differential and more favorable treatment to developing countries. The Enabling Clause also provides a legal basis for PTAs among developing countries and for the Global System of Trade Preferences (GSTP). Bhagwati (1991) has forcefully argued forcefully that PTAs are fundamentally incompatible with the WTO's stated goal of multilateral

¹Bhagwati and Panagariya (1996), Ray (1998), and Das and Ghosh (2006) contend that the majority of the PTAs have been formed between similar countries (so-called north-north agreements between developed countries and south-south agreements between developing countries) rather than between developed and developing countries (north-south agreements).

²As noted in Demir and Dahi (2011), South-South trade has become a substantial force in world trade. Between 1978 and 2005 the share of the South in world manufactures exports increased from 5% to 32%. The annual growth rate of real South-South manufactures exports has also been significantly higher than the world average reaching 14% as opposed to 6% for the latter. Krishna and Mitra (2008) also points out that by the late 1980s there was a "rush to free trade" among countries in the developing world (see Rodrik, 1994).

³While the WTO sanctions PTAs, it does so only under certain conditions. In particular, Article XXIV requires that (1) a PTA should cover "substantially all trade" between members; (2) a PTA should result in significant trade liberalization among members almost to the point that it leads to free trade amongst them; and (3) that PTA members not raise tariffs on non-members – a condition that appears to be an attempt to safeguard the interests of those left outside such discriminatory trade agreements. Of course, these conditions do not succeed in fully protecting the interests of non-members. It is also not necessarily the case that PTAs satisfying these requirements are consistent with multilateral trade liberalization.

trade liberalization and many researchers have questioned whether multilateral free trade is more likely to be achieved or the multilateral cooperation over free trade is more likely to be sustainable if the GATT allowed no exceptions to non-discriminatory trade liberalization.⁴

The main questions we address in this paper concern the role of PTAs for the multilateral negotiations over free trade. First, in a variation of the oligopoly trade model of Brander and Krugman (1983), we examine how different types of PTAs (North-North, South-South and North-South) affect the static welfare of member and non-member countries and the world as a whole, among countries of different production efficiency.⁵ Then, using a model of infinitely repeated interaction, we examine whether these arrangements make multilateral free trade more or less likely to be sustainable. For the most part we consider the role of single bilateral PTAs, but also consider the implications of multiple bilateral agreements for the likelihood of sustainability of multilateral free trade. As we show, there are significant differences in these implications depending on whether the PTA is a free trade agreement (FTA) or a customs union (CU). As in Saggi (2006), we compare the static noncooperative (no agreement) national welfare maximizing tariffs with those under various PTAs. To determine when multilateral free trade is easier or harder to sustain, we use the infinite repetition of this static game, where cooperation is required to be self-enforcing as in Riezman (1991), Bagwell and Staiger (1997a, 1997b, 1998, 1999), Bond et al. (2001). As in these past models, countries balance the current benefit of deviating from free trade against the future losses it suffers under the permanent trade war that results from its defection. In our setup, any initial agreement among countries (FTA or CU) is considered permanent in nature, so when countries deviate from free trade they revert back to their initial agreement (if any) in the trade war periods.

In the static framework, we show that North-North PTA leads to a larger reduction in the external tariffs of member countries on non-member countries (greater tariff complementarity effect), while a South-South PTA results in the smallest reduction. Comparing agreement types, FTAs result in larger external tariff reductions than CUs, which leads to non-member countries being better off under FTAs but worse off under CUs. While world welfare is always higher with a PTA than without, North-North PTAs yield the highest world welfare while South-South PTAs yield the lowest among agreements (due to substantial trade diversion), with stronger tariff complementarity under FTAs inducing higher welfare of FTAs relative to CUs of the same type (i.e. comparing a North-South FTA to a North-South CU).

The dynamic implications of PTAs on the sustainability of free trade lead to significantly different results relative to the ones under the static framework. Underlying these differences are two channels that we identify to help understand the intuition of the paper: the *membership effect* and the *cost effect*. The *membership effect* is defined as the impact of being a member of a PTA relative to being a non-member on the relative benefit of defecting from free trade to the cost of defecting, holding the costs of all countries the same. In this way, the membership effects isolates the role of

⁴Recent literature endogenizes the formation of PTAs and examines whether multilateral free trade is more likely to be stable when countries are (or not) able to form PTAs. See Das and Ghosh (2006), Goyal and Joshi (2006), Furusawa and Konishi (2007), Saggi and Yildiz (2010, 2013) for a detailed discussion.

⁵Similar to Das and Ghosh (2006), we employ an asymmetric four country model in order to accommodate the division of world economy into more than one trade bloc at a time.

PTA membership (that is, being a member or being a non-member) as if all countries in the model were symmetric. The *cost effect* instead isolates the role of cost differences among countries on the benefits and costs of defection from free trade, holding membership status constant. As the impact on outside (non-member) countries radically differs between FTAs and CUs, the membership effect works in opposite directions across the agreement types. On the other hand, the cost effect works in the same direction regardless of the agreement: high cost countries are less willing to cooperate over multilateral free trade. These two effects combine to result in a dynamic ranking of FTAs that is the opposite of the static welfare ranking when the production efficiency is moderately asymmetric: multilateral free trade is more likely to be sustainable under a South-South relative to all other agreements. Since the membership effect works differently for CUs, this reversal does not occur. We show that countries are most willing to cooperate over free trade under a North-North CU and least willing under a South-South CU. Finally, we find that with multiple bilateral agreements, global free trade is more likely when agreements are mixed (North-South) rather than polarized (North-North and South-South).

There are two key closely related papers to our line of research. In a similar static model, Das and Ghosh (2006) provides an endogenous FTA formation model that generates FTAs among similar countries. Unlike Das and Ghosh (2006), we examine both FTAs and CUs and focus on the role of the formation of PTAs for the sustainability of the multilateral cooperation over free trade. Whereas Das and Ghosh (2006) differentiates countries by market size, we employ differences in production cost to separate the North countries from the South countries. Saggi (2006) develops a symmetric three country model to examine similar dynamic questions. However, the present paper develops a four-country model in an asymmetric world economy. The advantage of having a four-country set-up, as opposed to the standard three-country model, is that it can accommodate the division of world economy into more than one trade bloc at a time.

The remainder of the paper is organized as follows. Section 2 presents the static tariff game between three two sets of asymmetric countries, two lower cost ‘North’ countries and two higher cost ‘South’ countries. Here we consider no agreement, a bilateral FTA and a bilateral CU and compare equilibrium tariff and welfare levels. The next section introduces the dynamic game, an infinite repetition of the static game and evaluates the incentives for countries to deviate from multilateral free trade under different PTAs relative to no agreement. In Section 4, we extend the model to incorporate multiple bilateral agreements and again investigate the incentives to deviate from global free trade. We offer our conclusions in Section 5.

2 The model

We develop a simple oligopoly model of trade between two regions: North and South. There are four countries: n , n' , s and s' . Countries n and n' are symmetric North countries while countries s and s' are symmetric South countries.⁶ The regions are distinguished by the asymmetry in production

⁶From hereon, for notational simplicity, n denotes the representative north firm and country while s denotes the representative south firm and country.

costs: $c_s > c_n$ where c_s (c_n) denotes the marginal cost of production of South (North) firms.⁷ Two goods are produced in each country: x and y . Good x is produced by a single profit-maximizing firm in each country at a constant marginal cost in terms of the numeraire good y that is produced under perfect competition with constant returns to scale.⁸ For simplicity, we use the same notation for firms as that of their corresponding countries.

Preferences over the two goods are assumed to be quasilinear:

$$U_i(x_i, y_i) = u(x_i) + y_i \quad (1)$$

where x_i is the total output sold in country i : $x_i \equiv \sum_j x_{ji}$ and x_{ji} denotes the output sold by country j 's firm in country i . Furthermore, $u(x_i)$ is assumed to be quadratic so that the demand curve for good x is linear in each country:

$$p_i(x_i) = \alpha - x_i \quad (2)$$

In the absence of any PTAs, firm j faces a specific tariff t_{ij} when exporting to country i . Since Article I of the GATT forbids tariff discrimination, we restrict attention to the case where $t_{ij} = t_i$ for all i and j .⁹ Firm i 's effective marginal cost of exporting to country z , denoted by c_{iz} , equals:

$$c_{iz} = c_i + t_z \text{ where } z \neq i \quad (3)$$

Firm j 's profits from exporting x_{ji} to country i , denoted by π_{ji} , can be written as:

$$\pi_{ji} = p_i(x_i)x_{ji} - c_{ji}x_{ji} \quad (4)$$

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

$$p_i + p'_i x_{ji} = c_{ji} \quad (5)$$

Utilizing the demand function in (2) and the above FOCs can be solved for equilibrium outputs and profits of all firms:

$$x_{ii} = \frac{\alpha - 4c_i + \sum_{z \neq i} c_z + 3t_i}{5}, x_{ki} = \frac{\alpha - 4c_k + \sum_{z \neq k} c_z - 2t_i}{5} \text{ and } x_i = \sum_z x_{zi} = \frac{4\alpha - \sum_z c_z - 3t_i}{5}$$

$$\pi_{ii} = x_{ii}^2 \text{ and } \pi_{ji} = x_{ji}^2 \quad (6)$$

The following comparative statics are standard:

$$\frac{\partial x_{zi}}{\partial t_i} < 0 < \frac{\partial x_{ii}}{\partial t_i}; \text{ and } \frac{dx_i}{dt_i} < 0 \text{ where } i \neq z \quad (7)$$

⁷For simplicity, fixed costs and transportation costs are assumed to be zero.

⁸The gains from trade stem from reduced market power in the domestic industry. To this end, the monopoly assumption is not crucial but is the simplest way to represent market power.

⁹It is obvious that $t_{ii} = 0$ for all i since countries do not impose taxes on their local firms and there are no tariffs on the numeraire good (that may be traded internationally in order to balance trade).

In other words, country i 's tariff lowers exports to its market (x_{zi}); increases the sales of its local firm (x_{ii}); and lowers the total output sold (x_i) in its market.

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

$$w_i(\mathbf{t}) \equiv S_i(t_i) + \sum_{z \neq i} \pi_{iz}(t_z) \quad (8)$$

where $\pi_{iz} = (p_i - c_{iz})x_{iz}$ denotes firm i 's export profit in country z and $\mathbf{t}$ is the global tariff vector $\mathbf{t} = (t_s, t_{s'}, t_n, t_{n'})$. Domestic surplus $S_i(t_i)$ is defined as:

$$S_i(t_i) \equiv u(x_i) - p_i x_i + t_i \sum_{z \neq i} x_{zi} + \pi_{ii} \quad (9)$$

where $u(x_i) - p_i x_i$ is consumer surplus in country i ; $\pi_{ii} = (p_i - c_i)x_{ii}$ equals firm i 's profits in its own market; and $t_i \sum_{z \neq i} x_{zi}$ is tariff revenue of country i .

3 Static tariff game

As in Brander and Krugman (1983), in our model, each country has a unilateral incentive to impose rent extracting tariffs on those trading partners with whom it does not have any trade agreement. Since markets are segmented and marginal costs are constant, strategic independence of trade policies obtains and own tariffs do not affect export profits. Let t_i^r denote the optimum tariff imposed by country i and $w_i(r)$ denote the equilibrium welfare of country i under trade regime r .

3.1 No agreement

Under no agreement $\{\Phi\}$, to maximize their welfare, all countries simultaneously choose their respective non-discriminatory (MFN) tariffs. Given the market segmentation, country i 's optimal MFN tariff choice problem reduces to the maximization of its domestic surplus:

$$\max_{t_i} S_i(t_i) \equiv u(x_i) - p_i x_i + t_i \sum_{z \neq i} x_{zi} + (p_i - c_i)x_{ii} \quad (10)$$

For simplicity, without loss of generality, we assume that the following holds hereafter: $c_s = c > c_n = 0$. The following optimal tariffs obtain under no agreement $\langle \Phi \rangle$:

$$t_n^\Phi = \frac{9\alpha - 2c}{33} \geq t_s^\Phi = \frac{9\alpha - 7c}{33} \quad (11)$$

3.2 Preferential trade agreements

Given that our objective is to examine the static and dynamic implications of different forms of PTAs, we allow single bilateral PTAs only.¹⁰ Therefore, beside $\{\Phi\}$, the following trade policy regimes can arise: (i) North-North PTA $\{nn'\}$; (ii) South-South PTA $\{ss'\}$; (iii) North-South PTA $\{ns\}$. In order to differentiate CUs from FTAs, we use a superscript u for CU, i.e. while $\{nn'^u\}$ denotes North-North CU, $\{nn'\}$ denotes North-North FTA.

¹⁰Later on, we will examine the implications of having additional bilateral PTAs.

3.3 Free trade agreements

Suppose countries i and j enter into an FTA $\{ij\}$ under which they impose zero tariffs on each other: $t_{ij} = t_{ji} = 0$. Under $\{ij\}$, member country i solves

$$\max_{t_i} W_i(t_i, \mathbf{t}_z), \text{ where } z \neq i, j \quad (12)$$

where $\mathbf{t}_z$ is a vector of tariffs member countries' firms face in non-member countries' markets while t_i denotes the external tariff country i imposes on the imports from non-member countries. The following optimal tariffs obtain:

$$t_n^{nn} = \frac{\alpha - 3c}{8}; t_s^{ss} = \frac{\alpha + 2c}{8}; t_n^{ns} = \frac{6\alpha + 7c}{48} \text{ and } t_s^{ns} = \frac{6\alpha - 13c}{48} \quad (13)$$

We show in the appendix that we have to impose a restriction on the degree of cost asymmetry in order to guarantee market access of south firms in export markets: $c \leq \bar{c}^f = \frac{3\alpha}{19}$.¹¹

3.4 Customs unions

As per Article XXIV of the GATT, when countries form a CU (say $\{ij^u\}$), we assume that they eliminate the internal tariffs on each other ($t_{ij} = t_{ji} = 0$) and impose common external tariff ($t_i(ij^u) = t_j(ij^u) = t^{ij^u}$) on the non-member countries. As a result, due to the market segmentation, member countries (i and j) solve the following problem:

$$\max_{t^{ij^u}} w_i(t^{ij^u}, \mathbf{t}_z) + w_j(t^{ij^u}, \mathbf{t}_z), \text{ where } z \neq i, j \quad (14)$$

where $\mathbf{t}_z$ is a vector of tariffs member countries' firms face in non-member countries' markets while t^{ij^u} denotes the common external tariff member countries impose on the imports from non-member countries. The following optimal tariffs obtain under CUs:

$$t^{nn^u} = \frac{\alpha - c}{4}; t^{ss^u} = \frac{\alpha}{4} \text{ and } t^{ns^u} = \frac{2\alpha - c}{8} \text{ where } z = n, s \quad (15)$$

where $z = s, n$. Note that export of south firms under all CU regimes is non-negative iff $c \leq \bar{c}^u = \frac{2\alpha}{21}$ holds. Since $\bar{c}^u < \bar{c}^f$ holds, in order to guarantee market access of south firms under any trade regime, we assume that $c < \bar{c}^u$ holds hereafter.

In order to minimize the potential harmful effects of PTAs, Article XXIV of the GATT requires that member countries not raise tariffs (or any other trade restrictions) on non-members. In our model as in Bagwell and Staiger (1997a, 1999), we find that this requirement is not binding since the formation of a bilateral PTA induces each member to lower its tariff on the non-member country relative to the status quo (i.e. the model exhibits tariff complementarity).¹² The intuition can be explained as follows: when PTA members reduce their tariff on each other to zero, exports of the

¹¹The complete derivation and formulae we do not report in the text can be found in the appendix.

¹²World Bank (2005) argues that no clear evidence shows that the formation of a PTA leads member countries to become more protectionist towards non-member countries. Magee and Lee (2001) shows that the formation of European Economic Community (EEC) induced members to reduce their external tariffs. More recently, Limao (2006), Karacaovali and Limao (2008) and Estevadeordal et al. (2008) empirically show that the tariff complementarity effect arises under FTAs.

non-member countries decrease and thus they become a less important rent-extracting source. We denote the tariff complementarity effect realized under an FTA $\{ij\}$ and under a CU $\{ij^u\}$ in a member country i by $\psi_i(ij) = t_i^\Phi - t_i^{ij} > 0$ and $\psi_i(ij^u) = t_i^\Phi - t_i^{j^u} > 0$, respectively. We can directly state the following result that is immediate from the tariffs reported above:

Proposition 1 *The tariff complementarity effect is larger under an FTA relative to a CU of the same type: $\psi_i(ij) > \psi_i(ij^u)$ and the formation of a North – North (South – South) PTA leads to the largest (smallest) external tariff reduction relative to no agreement.*

The intuition behind the first part of the above proposition is as follows. Unlike an FTA, a CU takes into account the export profits that members derive from each other's markets and thus members impose higher external tariffs on non-members under a CU relative to an FTA of the same type. The second part of the proposition argues that the *tariff complementarity effect* decreases (increases) in the average cost of the member (non-member) countries. Intuitively, when the member countries of a PTA are more efficient than the non-member countries, the non-members' exports to member countries fall relatively more, further reducing the protection incentives of the member countries. Consequently, since the average cost of the member countries is the lowest, a North-North PTA yields the largest tariff complementarity effect while the opposite is true for a South-South PTA. Next, we focus on welfare:

Proposition 2 *While member countries are better off under any PTA relative to no agreement, non-member countries are better off under FTAs and worse off under CUs relative to no agreement. Aggregate world welfare under a PTA is always larger relative to that under no agreement while North-North PTA (South-South PTA) yields the highest (lowest) world welfare.*

From each country's perspective, a PTA embodies the following trade-off. On the one hand, forming a PTA lowers a member country's domestic surplus relative to the case where it can use its optimally chosen tariff(s). On the other hand, being part of a PTA increases export profits in other member's market. We show in the appendix that, when $c \leq \bar{c}^u$ holds, the latter effect dominates the former so that countries always have an incentive to form a PTA. Furthermore, we can make two immediate points for non-member countries: (i) due to market segmentation, a PTA does not alter the domestic surplus of a non-member country and (ii) quantity competition in the product market implies that equilibrium export profits are directly related with the total exports. Since the non-member countries' exports increase when the other countries form an FTA, their welfare levels increase relative to the no agreement. On the other hand, non-member countries' exports decrease when the other countries form a CU as the tariff complementarity is not large enough to offset the discrimination they face relative to member countries and their welfare decreases relative to no agreement.¹³ This result implies that the PTAs can have harmful effects on the non-member countries unless they are accompanied by sufficient external tariff reductions. The second part of the above proposition argues that, since North-North PTA leads to larger trade creation and smaller trade diversion relative to a South-South PTA (and North-South PTA), world welfare follows the same ranking as that of the tariff complementarity effects.

¹³Winters (1997) and Chang and Winters (2002) argue that welfare of non-member countries need not be perfectly correlated with their exports. Our results mainly rely on the assumptions of market segmentation and Cournot competition.

The above proposition does not speak about the welfare of *member* countries under FTAs relative to CUs of the same type. When a PTA involves symmetric countries, member countries are better off under a CU relative to an FTA of the same type. This stems from the fact that, unlike an FTA, a CU takes into account the export profits that members derive from each others' markets in joint welfare maximization. Thus, we obtain

$$w_n(nn'^u) > w_n(nn') \text{ and } w_s(ss'^u) > w_s(ss') \text{ for all } c \quad (16)$$

However, when a PTA involves asymmetric members (such as North-South PTA), while common external tariff setting always benefits the north member, it harms the south member country if it is sufficiently high cost:

$$w_n(ns^u) > w_n(ns) \text{ for all } c \text{ while } w_s(ns^u) > w_s(ns) \text{ iff } c < c_s^{ns} \quad (17)$$

where $c_s^{ns} < \bar{c}^u$ holds.

So far, since we have not allowed for multilateral tariff reductions, the only way to achieve tariff levels lower than the one under the no agreement is via bilateral PTAs. Next we examine the following question: how does bilateral trade liberalization (via an FTA or a CU) affect incentives for multilateral tariff cooperation over free trade?

4 Multilateral cooperation over free trade

Similar to the existing literature, multilateral tariff cooperation over free trade is modeled as a stationary repeated game where cooperation can be sustained only if it is incentive compatible for all countries.¹⁴ In previous models, cooperation is self-enforcing in the sense that each country balances the current benefit of deviating from free trade against the future losses caused by the breakdown of multilateral cooperation that results from its defection. We assume that countries sustain cooperation via trigger strategies and defection by any country results in a multilateral trade war. It is useful to note an important difference between no agreement and PTAs. By assumption, PTAs are permanent in nature so that member countries of a PTA retain zero tariffs on each other even if cooperation with non-member countries breaks down. On the other hand, countries revert to optimally chosen individual MFN tariffs in the absence of PTAs.

4.1 Cooperation under no agreement

Suppose that each country employs zero tariff until any one of them defects. Once defection happens, countries switch back to their optimally chosen individual MFN tariffs. Let $W_i^\Phi(\mathbf{t} = 0)$ denote the per period welfare of country i under cooperation:

$$W_i^\Phi(\mathbf{t} = 0) = S_i^\Phi(t_i = 0) + \sum_{z \neq i} \pi_{iz}(t_z = 0) \quad (18)$$

¹⁴See Riezman (1991), Bagwell and Staiger (1997a, 1997b, 1998, 1999), Bond et al. (2001), and Saggi (2006).

where $S_i^\Phi(t_i = 0)$ is the domestic surplus and $\sum_{z \neq i} \pi_{iz}(t_z = 0)$ equals the total exporting profit of country i under free trade. Next, we define the welfare of country i that defects from zero tariff to its optimal MFN tariff under no agreement (t_i^Φ) while other countries cooperate:

$$W_i^\Phi(t_i^\Phi) = S_i^\Phi(t_i^\Phi) + \sum_{z \neq i} \pi_{iz}(t_z = 0) \quad (19)$$

Note that since markets are strategically separated from each other, defection from free trade to t_i^Φ does not affect the total export profits of country i whereas it increases its domestic surplus from $S_i^\Phi(t_i = 0)$ to $S_i^\Phi(t_i^\Phi)$. Since defection by any country leads to a permanent multilateral trade war where all countries impose their optimally chosen individual MFN tariffs, welfare levels under trade war (No agreement) are as follows:

$$W_i^\Phi(\mathbf{t}^\Phi) = S_i^\Phi(t_i^\Phi) + \sum_{z \neq i} \pi_{iz}(\mathbf{t}_z^\Phi) \quad (20)$$

where $\mathbf{t}^\Phi$ denotes the vector of optimally chosen tariffs by all countries while $\mathbf{t}_z^\Phi$ excludes t_i^Φ from $\mathbf{t}^\Phi$. Using the expressions in (18), (19) and (20), we are able to define one period benefit and cost of defection for countries. First, we define the benefit from the defection:

$$\begin{aligned} B_i^\Phi(t_i^\Phi) &= W_i^\Phi(t_i^\Phi) - W_i^\Phi(\mathbf{t} = 0) \\ &= \underbrace{S_i^\Phi(t_i^\Phi) - S_i^\Phi(t_i = 0)}_{>0} \end{aligned} \quad (21)$$

Similarly, per period cost of defection to a country is given by¹⁵

$$\begin{aligned} C_i^\Phi(\mathbf{t}^\Phi) &= W_i^\Phi(\mathbf{t} = 0) - W_i^\Phi(\mathbf{t}^\Phi) \\ &= \underbrace{[S_i^\Phi(t_i = 0) - S_i^\Phi(t_i^\Phi)]}_{-B_i^\Phi(t_i^\Phi, t) < 0} + \underbrace{\left[\sum_{z \neq i} \Pi_{iz}(t_z = 0) - \sum_{z \neq i} \Pi_{iz}(\mathbf{t}_z^\Phi) \right]}_{>0} \end{aligned} \quad (22)$$

In order for tariff cooperation to be self-enforcing, the one period benefit from defection must be less than the discounted lifetime cost of defection since defection leads to permanent multilateral trade war. In other words, the incentive compatibility (IC) constraint must hold for each country as follows:

$$B_i^\Phi(t_i^\Phi) \leq \frac{\delta}{1-\delta} C_i^\Phi(\mathbf{t}^\Phi)$$

where δ denotes the discount factor and $\frac{\delta}{1-\delta} C_i^\Phi(\mathbf{t}^\Phi)$ measures the trade war's life-time cost of defection to each country under $\{\Phi\}$. For each country, the critical discount factor δ_i^Φ above which cooperation over free trade is self-enforcing obtains when $B_i^\Phi(t_i^\Phi) = \frac{\delta}{1-\delta} C_i^\Phi(\mathbf{t}^\Phi)$ holds.

¹⁵Note that under complete symmetry ($c_s = 0$), following Brander and Spencer (1984), it is straightforward to argue that the second term in (22) is greater than the absolute value of the first term so that $C_i^\Phi(\mathbf{t}^\Phi, t)$ is always positive. This implies that under symmetry, if countries were to completely discount the future payoffs, multilateral tariff cooperation would not be feasible.

Next, we examine the nature of per period benefit and cost of defection from the perspective of South and North countries. Before proceeding, it is important to note that countries' incentives for cooperation differ across the regions because of the underlying differences in production costs. To this end, the following lemma summarizes how the benefit and cost of defection from tariff cooperation under no agreement changes with the own and other countries' costs of production.

Lemma 1 *Let $z \neq i$. Then, (i) $\frac{\partial B_i^\Phi(t_i^\Phi)}{\partial c_i} < 0$ and $\frac{\partial B_i^\Phi(t_i^\Phi)}{\partial c_z} < 0$; (ii) $\frac{\partial C_i^\Phi(\mathbf{t}^\Phi)}{\partial c_i} < 0$ and $\frac{\partial C_i^\Phi(\mathbf{t}^\Phi)}{\partial c_z} > 0$.*

The intuition behind the first part of the above lemma has to do with the fact that tariffs are used to extract rents in the model: since demand is symmetric across countries, any given reduction in imports decreases consumer welfare in a roughly equal way in all countries whereas it results in a greater increase in domestic firm's local profits in lower cost countries due to higher mark-ups. Thus, the benefit from defection rises as a country becomes more efficient. Similarly, when the rival countries become more efficient, the incentive to extract rent increases and thus the benefit from defection rises.

The intuition for the second part can be explained as follows. Due to the smaller volume of exports, countries with higher costs of production benefit less from tariff reductions granted by others. Similarly, such countries have relatively more to lose from eliminating their own optimal tariffs since they apply to relatively larger import volumes. Thus, the cost of defection from tariff cooperation of a country depends negatively on its own cost. Finally, the higher the cost of the other countries, the larger the increase in export profits due to the tariff cooperation over free trade, and the smaller the loss due to its own trade liberalization since tariff reduction applies to smaller volume of imports (due to higher cost of the other countries). Therefore, the cost of defection falls when other countries become more efficient. Since the benefit from defection rises while the cost of defection falls when the other countries become more efficient, the following proposition is immediate:

Proposition 3 *Under $\{\Phi\}$, the range of discount factors above which north country is willing to cooperate over free trade is larger than that above which south countries are willing to cooperate: $\delta_s^\phi \geq \delta_n^\phi$. Thus, multilateral cooperation over free trade is sustainable if and only if $\delta \geq \delta_s^\phi$.*

[Insert Figure 1 here]

The above proposition informs us that the higher the production cost of a country, the less willing it is to cooperate multilaterally over free trade. Therefore, under no agreement, the critical discount factor above which south countries are willing to cooperate over free trade binds for the sustainability of multilateral cooperation over free trade.

Next, consider how the formation of PTAs of different types changes the incentives for multilateral cooperation over tariffs.

4.2 Cooperation under PTAs

Suppose two countries sign a binding PTA and eliminate tariffs on each other. As mentioned above, we assume that countries are bound to their PTAs, implying that member countries retain zero

tariffs on each other even when the multilateral cooperation breaks down. How does the formation of a PTA $\{ij\}$ if it is an FTA or $\{ij\}^u$ if it is a CU) alter the costs and benefits of multilateral tariff cooperation for member and nonmember countries? To address this question, first consider the perspective of the non-member countries.

4.2.1 Cooperation incentives of non-members under PTAs

It is straightforward to argue that when countries cooperate over free trade, the per period welfare of a nonmember country under free trade and in the defection period (thus the benefit of defection from cooperation for the non-member country) stays the same under a PTA as under $\{\Phi\}$. After the breakdown of cooperation, a non-member country (say country k) responds via raising its tariff from zero to t_k^Φ as under $\{\Phi\}$ from next period on, but faces t_z^Φ in the other non-member country z and t_i^{ij} and t_j^{ij} in member countries' markets under an FTA ($t_i^{ij^u}$ and $t_j^{ij^u}$ under a CU). Since the benefit from the defection of a non-member country under a PTA is the same as the one under no agreement, it behaves the same way as outlined in Lemma 1: $\frac{\partial B_k^r(t_k^\Phi)}{\partial c_k} < 0$ and $\frac{\partial B_k^r(t_k^\Phi)}{\partial c_z} < 0$, where $z \neq k$ and $r = \{ij\}, \{ij\}^u$. However, for a non-member country, the formation of a PTA yields a reinforcing effect on the cost of defection. The lower the cost of the member countries, the larger the tariff complementarity effect under a PTA the non-member country enjoys. Thus, cost of defection from cooperation falls as member countries become more efficient: $\frac{\partial C_k^r(t_i^r, t_j^r, t_k^\Phi)}{\partial c_i} > \frac{\partial C_k^\Phi(\mathbf{t}^\Phi)}{\partial c_i} > 0$, where $r = \{ij\}, \{ij\}^u$. Then the following ranking is immediate:

$$\begin{aligned} \delta_s^{nn'} &\geq \delta_{s'}^{ns} \geq \delta_{n'}^{ns} > \delta_n^{ss'} \\ \delta_s^{nn'^u} &\geq \delta_{s'}^{ns^u} \geq \delta_{n'}^{ns^u} > \delta_n^{ss'^u} \end{aligned} \quad (23)$$

The above ranking provides support for the idea that the lower the cost of member countries of a PTA, the less willing the non-member countries to cooperate over free trade. Similarly, when non-member countries are asymmetric (as under a North-South PTA), the south non-member country's incentive to cooperate is lower than that of the north non-member country. Note that the critical discount factors for the non-member countries follow the same ranking of the tariff complementarity effects of PTAs (see Proposition 1).

Next, we compare the incentives of non-member countries under a PTA relative to no agreement. It is immediate from the second part of the proposition 2 that a non-member country is better (worse) off under an FTA (a CU) relative to no agreement. Thus, the following is immediate:

Proposition 4 *The per period cost to a non-member country of the breakdown of cooperation is smaller (greater) under an FTA (a CU) relative to that under $\{\Phi\}$ while the benefit of defection stays the same. Therefore, an FTA (a CU) makes non-member countries less (more) willing to cooperate multilaterally over free trade: (i) $\delta_s^{nn'} > \delta_s^\Phi > \delta_{s'}^{nn'^u}$; (ii) $\delta_{s'}^{ns} > \delta_s^\Phi > \delta_{s'}^{ns^u}$; (iii) $\delta_{n'}^{ns} > \delta_n^\Phi > \delta_{n'}^{ns^u}$ and (iv) $\delta_n^{ss'} > \delta_n^\Phi > \delta_n^{ss'^u}$.*

Combining the ranking in (23) and Proposition 4, we can summarize the ranking structure of critical discount factors: $\delta_s^{nn'} \geq \delta_{s'}^{ns} \geq \delta_s^\Phi$ and $\delta_s^\Phi > \delta_{s'}^{nn'^u} \geq \delta_{s'}^{ns^u}$. The results so far speak out for the sustainability of multilateral cooperation over free trade only from non-member countries' perspectives. Since the incentive constraint of the least willing participant determines the sustainability

of multilateral free trade, we can argue by focusing on non-member countries that if multilateral free trade cannot be supported under no agreement ($\delta < \delta_s^\phi$), it can never be sustained under the North-North FTA and North-South FTA as well (the formation of these agreements hinders multilateral tariff cooperation). Unlike bilateral FTAs, non-member countries under CUs of the same type have stronger incentives to cooperate relative to no agreement. In order to see the full picture, we have to consider the incentives of member countries for multilateral tariff cooperation under PTAs. To this end, we examine FTAs and CUs in isolation.

4.2.2 Cooperation incentives of FTA members

Under multilateral tariff cooperation over free trade, the per period welfare of a member country (say i) under $\{ij\}$ equals

$$W_i^{ij}(\mathbf{t} = 0) = S_i^{ij}(t_i = 0) + \sum_{z \neq i} \pi_{iz}(t_z = 0) \quad (24)$$

Since an FTA is permanent by assumption, in the defection period and under permanent trade war thereafter, the FTA member i raises its tariff (from zero to t_i^{ij}) only on non-member countries. Thus, a member country's welfare in the period of defection equals

$$W_i^{ij}(t_i^{ij}) = S_i^{ij}(t_i^{ij}) + \sum_{z \neq i} \pi_{iz}(t_z = 0) \quad (25)$$

while the per period welfare of a member country under a trade war equals

$$W_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) = S_i^{ij}(t_i^{ij}) + \pi_{ij}(t_j^{ij}) + \sum_{z \neq i, j} \pi_{iz}(t_z^\Phi), \text{ where } z \neq i, j \quad (26)$$

An FTA member's benefit from defection equals the increase in domestic surplus due to the ability to raise its tariff from zero to t_i^{ij} :

$$\begin{aligned} B_i^{ij}(t_i^{ij}) &= W_i^{ij}(t_i^{ij}) - W_i^{ij}(\mathbf{t} = 0) \\ &= \underbrace{S_i^{ij}(t_i^{ij}) - S_i^{ij}(t_i = 0)}_{>0} \end{aligned} \quad (27)$$

When a member country i defects from cooperation, from next period on, it faces t_z^Φ in non-members' markets while the other FTA member j raises its tariff on non-members from zero to t_j^{ij} . Then, a member's per period cost of defection equals

$$\begin{aligned} C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) &= W_i^{ij}(\mathbf{t} = 0) - W_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) \\ &= \underbrace{S_i^{ij}(t_i = 0) - S_i^{ij}(t_i^{ij})}_{-B_i^{ij}(t_i^{ij}) < 0} + \sum_{z \neq i} \pi_{iz}(t_z = 0) - \pi_{ij}(t_j^{ij}) - \sum_{z \neq i, j} \pi_{iz}(t_z^\Phi) \end{aligned} \quad (28)$$

In order to examine a member country's benefit and cost of defection from multilateral cooperation over free trade under an FTA $\langle \{ij\} \rangle$, we distinguish two distinct effects. On one hand, relative

to no agreement, the inability of a member country i to raise its tariff on its FTA partner j reduces the benefit from defection. Then, it is immediate to argue from (28) that the first component of the cost of defection ($-B_i^{ij}(t_i^{ij})$) increases. On the other hand, in the punishment phase (under a permanent trade war) while all countries suffer from the tariff increase under no agreement, a member country under an FTA partially benefits due to the tariff increase in its partner's market since it applies only to non-member countries. The following lemma summarizes how the formation of an FTA alters the costs and benefits of multilateral tariff cooperation for member countries.

Lemma 2 *Let $z \neq i, j$ denotes the non-member country under an FTA $\langle\{ij\}\rangle$. Then, the following obtains: (i) $\frac{\partial B_i^{ij}(t_i^{ij})}{\partial c_i} < 0$, $\frac{\partial B_i^{ij}(t_i^{ij})}{\partial c_j} > 0$ and $\frac{\partial B_i^{ij}(t_i^{ij})}{\partial c_z} < 0$ and (ii) $\frac{\partial C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi)}{\partial c_i} < 0$, $\frac{\partial C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi)}{\partial c_j} < 0$ and $\frac{\partial C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi)}{\partial c_z} > 0$.*

The above lemma implies two noteworthy results: (i) since $\frac{\partial B_i^{ij}(t_i^{ij})}{\partial c_j} > 0$ and $\frac{\partial C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi)}{\partial c_j} < 0$ hold, if member countries are asymmetric, the one with a higher cost is less willing to cooperate relative to its FTA partner (with a lower cost) and (ii) since $\frac{\partial B_i^{ij}(t_i^{ij})}{\partial c_z} < 0$ and $\frac{\partial C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi)}{\partial c_z} > 0$ hold, member countries are more willing to cooperate as the non-member countries become less efficient. Therefore, we obtain the following ranking:

$$\delta_s^{ss'} \geq \delta_s^{ns} \geq \delta_n^{ns} \geq \delta_n^{nn'} \quad (29)$$

4.3 Sustainability of multilateral free trade under FTAs

Sustainability hinges on the discount factor of the country that is least willing remain in the agreement. To this end, we identify two distinct effects that arise upon the formation of an FTA: the *membership effect* and the *cost effect*. First, we isolate the former effect by assuming that all countries are symmetric with respect to cost of production ($c = 0$). Under such a case, while both the benefit and cost of defection for members are lower (countries i and j) relative to that of non-members, the difference in the cost of defection is relatively smaller:

$$B_i^{ij}(t_i^{ij}) - B_z^{ij}(t_z^\Phi) = -\frac{3\alpha^2}{88} < 0, \text{ where } z \neq i, j \quad (30)$$

$$C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) - C_z^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) = -\frac{21\alpha^2}{3872} < 0, \text{ where } z \neq i, j \quad (31)$$

and

$$B_z^{ij}(t_z^\Phi) - B_i^{ij}(t_i^{ij}) > C_z^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi) - C_i^{ij}(t_i^{ij}, t_j^{ij}, t_z^\Phi), \text{ where } z \neq i, j \quad (32)$$

Thus, when countries are symmetric, non-member countries are less willing to cooperate relative to member countries: $\delta_z^{ij} > \delta_i^{ij} = \delta_j^{ij}$, where $z \neq i, j$ and this effect is called as the *membership effect*. We also know from Lemma 1 and Lemma 2 (and the subsequent discussions) that the higher the cost of a country the less willing it is to multilaterally cooperate over free trade and this is called the *cost effect*, which becomes stronger as the cost asymmetry gets larger. As represented in figure 2a, since these two effects reinforce each other under $\{nn'\}$, the south non-member countries' incentive is binding for the sustainability of multilateral cooperation.

Proposition 5a *Under $\{nn'\}$, multilateral cooperation over free trade is sustainable if and only if $\delta \geq \delta_s^{nn'}$.*

[Insert Figure 2a]

Similarly, under North-South agreement $\{ns\}$, both the *membership effect* and the *cost effect* work in the same direction for the south non-member country which, as the least willing country, determines the critical discount factor. Here it is worth mentioning that *cost effect* dominates the *membership effect* for the south member country relative to the north non-member country and thus we obtain (see Figure 2b):

$$\delta_{s'}^{ns} > \delta_s^{ns} > \delta_{n'}^{ns} > \delta_n^{ns} \quad (33)$$

The following is immediate:

Proposition 5b *Under $\{ns\}$, multilateral cooperation over free trade is sustainable if and only if $\delta \geq \delta_{s'}^{ns}$.*

[Insert Figure 2b]

Finally, consider South-South agreement $\{ss'\}$. We find that cost effect dominates the membership effect if and only if south countries are sufficiently high cost: $\delta_s^{ss'} \geq \delta_n^{ss'}$ if and only if $c \geq c^{ss'}$ (see Figure 2c). In other words, when countries are relatively symmetric, the north non-member country's incentive is pivotal for multilateral cooperation but the south member countries' incentives determine the critical discount factor when they are sufficiently high cost. Thus, we obtain the following result:

Proposition 5c *Under $\{ss'\}$, multilateral cooperation over free trade is sustainable if and only if $\delta > \max(\delta_s^{ss'}, \delta_n^{ss'})$ where $\delta_s^{ss'} \geq \delta_n^{ss'}$ iff $c \geq c^{ss'}$.*

[Insert Figure 2c]

Next, we compare the range of discount factors above which multilateral cooperation over free trade is sustainable under different types of FTAs. Before proceeding, let $c(r - v)$ denote the cost threshold at which the critical discount factors under FTAs r and v are equal. Combining (23), Proposition 5a, Proposition 5b, and Proposition 5c, we can state our main result in the FTA context (see Figure 5):

Proposition 6 *While the range of discount factors above which multilateral cooperation over free trade is sustainable is larger under $\{ns\}$ relative to $\{nn'\}$ for all c : $\delta_s^{nn'} \geq \delta_{s'}^{ns}$, we obtain $\max(\delta_s^{ss'}, \delta_n^{ss'}) \geq \delta_{s'}^{ns}$ iff $c \geq c(ns - ss')$ and $\max(\delta_s^{ss'}, \delta_n^{ss'}) \geq \delta_n^{nn'}$ iff $c \geq c(nn' - ss')$, where $c(ns - ss') < c(nn' - ss')$.*

[Insert Figure 3]

The first part of the above proposition is immediate from (23) and it argues in favor of the idea that while aggregate world welfare is higher under $\{nn'\}$ relative to $\{ns\}$ statically, multilateral free trade is easier to sustain under $\{ns\}$ relative to $\{nn'\}$. The primary reason for this result is that the non-member south country is less willing to cooperate as average efficiency of FTA members rises since the tariff complementarity effect (thus the free rider effect) is larger under $\{nn'\}$ relative

to $\{ns\}$. Thus, the external trade liberalization of FTA members itself reduces the incentives of non-members for multilateral trade liberalization making multilateral cooperation less likely to be sustainable.

The second part of the above proposition informs us that when south countries are sufficiently low cost ($c < c(ns - ss')$), multilateral free trade is easier to sustain under $\{ss'\}$ relative to other types of FTAs. To understand the intuition, note that since $c^{ss'} < c(ns - ss')$ holds we have to consider two cost ranges: (i) $\delta_s^{ss'} < \delta_n^{ss'}$ when $c < c^{ss'}$ and (ii) $\delta_s^{ss'} > \delta_n^{ss'}$ when $c(ns - ss') > c > c^{ss'}$. Suppose first that $c < c^{ss'}$ holds and $\delta_n^{ss'}$ is the critical discount factor under $\{ss'\}$. It is immediate from (23) that $\delta_n^{ss'} < \delta_{s'}^{ns}$ (thus $\delta_n^{ss'} < \delta_s^{nn'}$) holds since (i) country n is more efficient relative to country s' and (ii) average cost of members is lower under $\{ss'\}$ relative to $\{ns\}$. Consider now that $c(ns - ss') > c > c^{ss'}$ holds and $\delta_s^{ss'}$ is the critical discount factor under $\{ss'\}$. Over this range, both the benefit and cost of defection of country s under $\{ss'\}$ is lower than the ones of country s' under $\{ns\}$ and the difference in costs is smaller relative to the difference in benefits. Thus, cooperation is easier to sustain under $\{ss'\}$ relative to $\{ns\}$ and (thus relative to $\{nn'\}$). On the other hand, as c rises, the south member countries' cost of defection under $\{ss'\}$ falls more rapidly relative to the south non-member country under $\{ns\}$ and $\{nn'\}$ and thus we obtain that the critical discount factor under $\{ss'\}$ exceeds the one under $\{ns\}$ when c is sufficiently high: $c \geq c(ns - ss')$. Similar argument applies for the comparison of $\delta_s^{ss'}$ and $\delta_s^{nn'}$. As a result, when countries are relatively symmetric, multilateral cooperation over free trade is easier to sustain under $\{ss'\}$ relative to other types of FTAs. Under such a case, it is important to note that the critical discount factors under different FTAs follow the same ranking as of the aggregate world welfare. In other words, positive dynamic implications comes at the expense of a static welfare loss. However, when south countries are sufficiently high cost, the opposite result obtains and the static and dynamic implications become aligned.

4.3.1 Comparison of FTAs with no Agreement

So far we have examined the sustainability of multilateral cooperation over free trade under different types of FTAs. Next, we ask whether the formation of FTAs hinders or facilitates multilateral cooperation relative to no agreement. We know from Proposition 4 and Proposition 6 that $\{ns\}$ and $\{nn'\}$ raise the critical discount factor above which multilateral cooperation over free trade is sustainable relative to no agreement: $\delta_s^{nn'} \geq \delta_{s'}^{ns} > \delta_s^\phi$. What remains is the comparison of δ_s^ϕ and $\max(\delta_s^{ss'}, \delta_n^{ss'})$.

Proposition 7 *While multilateral cooperation over free trade is easier to sustain under $\{\Phi\}$ relative to $\{ns\}$ and $\{nn'\}$ for all cost levels, it is harder to sustain under $\{\Phi\}$ relative to $\{ss'\}$ when $\underline{c(\Phi - ss')} \geq c \geq \underline{c(\Phi - ss')}$ holds.*

[Insert Figure 4]

There has been widespread concern regarding that the formation of South-South trade agreements undermines multilateral trade liberalization efforts. However, the above proposition informs us that South-South FTA is the *only* bilateral FTA that can facilitate multilateral cooperation

when the degree of cost asymmetry across countries is at a medium range (see Figure 4). It is crucial to note that when south countries are sufficiently high cost multilateral free trade is less likely to be sustainable under South-South FTA relative to all other possible trade regimes.

We next consider the incentives of CU members for multilateral cooperation.

4.4 Cooperation incentives of CU members

As before, under multilateral tariff cooperation over free trade, the per period welfare of a member country (say i) under $\{ij^u\}$ equals

$$W_i^{ij^u}(\mathbf{t} = 0) = S_i^{ij^u}(t_i = 0) + \sum_{z \neq i} \pi_{iz}(t_z = 0) \quad (34)$$

Like FTAs, a CU (say $\{ij^u\}$) is assumed to be a permanent agreement so that both in the defection period and under permanent trade war members impose zero tariffs on each other and t^{ij^u} on the non-member countries. Here, it is important to note that, unlike an FTA, the defection of CUs from multilateral cooperation involves both members due to the requirement of common external tariff under CUs. Thus, a member country's welfare in the period of defection equals

$$W_i^{ij^u}(t^{ij^u}) = S_i^{ij^u}(t^{ij^u}) + \pi_{ij}(t^{ij^u}) + \sum_{z \neq i,j} \pi_{iz}(t_z = 0) \quad (35)$$

while the per period welfare of a member country under a trade war equals

$$W_i^{ij^u}(t^{ij^u}, t_z^\Phi) = S_i^{ij^u}(t^{ij^u}) + \pi_{ij}(t^{ij^u}) + \sum_{z \neq i,j} \pi_{iz}(t_z^\Phi) \quad (36)$$

A CU member's benefit from defection equals the increase in (i) domestic surplus due to the ability to raise its tariff from zero to $t_i^{ij^u}$ and (ii) export profits in CU partner's market since t^{ij^u} applies to non-member countries only:

$$\begin{aligned} B_i^{ij^u}(t^{ij^u}) &= W_i^{ij^u}(t^{ij^u}) - W_i^{ij^u}(\mathbf{t} = 0) \\ &= \underbrace{S_i^{ij^u}(t^{ij^u}) - S_i^{ij^u}(t_i = 0)}_{>0} + \underbrace{\pi_{ij}(t_j = t^{ij^u}) - \pi_{ij}(t_j = 0)}_{>0} \end{aligned} \quad (37)$$

Note that the second effect does not arise under an FTA since the defection of CUs from multilateral cooperation involves both members unlike FTAs. Thus, a member's benefit from defection under a CU is always larger relative to that under an FTA of the same type:

$$B_i^{ij^u}(t^{ij^u}) > B_i^{ij}(t_i^{ij}) \text{ for all } c$$

When a member i defects from cooperation, from next period on, it faces t_z^Φ in non-members' markets while the CU's tariff on non-members stays the same as t^{ij^u} . Then, a CU member's per period cost of defection equals

$$\begin{aligned} C_i^{ij^u}(t^{ij^u}, t_z^\Phi) &= W_i^{ij^u}(\mathbf{t} = 0) - W_i^{ij^u}(t^{ij^u}, t_z^\Phi) \\ &= -B_i^{ij^u}(t^{ij^u}) + \sum_{z \neq i,j} \pi_{iz}(t_z = 0) - \pi_{iz}(t_z^\Phi) \end{aligned} \quad (38)$$

It is immediate from (16) that cost of defection of member countries under $\{nn'^u\}$ and $\{ss'^u\}$ is smaller relative to the one under $\{nn'\}$ and $\{ss'\}$. Moreover, we know from (17) that while the north member's cost of defection under $\{ns^u\}$ is always lower relative to $\{ns\}$, the opposite obtains for the south member only when it is sufficiently high cost. The following proposition states that, over that range, the benefit effect dominates the cost effect for the south member country under $\{ns^u\}$ relative to $\{ns\}$:

Proposition 8 *Member countries have weaker incentive to cooperate multilaterally over free trade under a CU relative to an FTA of the same type: (i) $\delta_n^{nn'^u} > \delta_n^{nn'}$; (ii) $\delta_s^{ss'^u} > \delta_s^{ss'}$; (iii) $\delta_n^{ns^u} > \delta_n^{ns}$ and (iv) $\delta_s^{ns^u} > \delta_s^{ns}$.*

The above proposition along with Proposition 4 argues in favor of the idea that the range of discount factors for which a member (non-member) country is willing to cooperate over multilateral free trade is smaller (larger) under a CU than that under an FTA of the same type. Next, we examine the critical discount factors under different types of CUs

4.5 Sustainability of multilateral free trade under CUs

As under FTAs, we focus on the least willing country's incentive that determines the critical discount factor over which multilateral tariff cooperation is sustainable under a CU. Thus, we combine the incentives of non-member countries and member countries. Before proceeding, note that while the *membership effect* and the *cost effect* arise under the formation of both FTAs and CUs, the nature of these effects is quite different.

On one hand, the two factors together imply that the direction of the membership effect under a CU is the opposite to that under an FTA: (i) the common external tariff under a CU is higher than the individual external tariff under an FTA and (ii) unlike an FTA, the defection of CUs from multilateral cooperation involves both members. As before, in order to isolate the membership effect, we assume that cost effect does not exist ($c = 0$). We show that, when countries are completely symmetric, the benefit from defection for members (countries i and j) is higher while their cost of defection is lower relative to non-members:

$$B_i^{ij^u}(t^{ij^u}) - B_z^{ij^u}(t_z^\Phi) = \frac{1}{11}\left(\frac{\alpha}{10}\right)^2 > 0$$

and

$$C_i^{ij^u}(t^{ij^u}, t_z^\Phi) - C_z^{ij^u}(t^{ij^u}, t_z^\Phi) = -\frac{1069}{2}\left(\frac{\alpha}{110}\right)^2 < 0$$

Thus, opposite to FTAs, member countries are less willing to cooperate relative to non-member countries under complete symmetry: $\delta_z^{ij} < \delta_i^{ij} = \delta_j^{ij}$, where $z \neq i, j$ and $c = 0$. In other words, the membership effect acts in the reverse direction under the formation of a CU relative to an FTA.

On the other hand, the joint determination of external tariffs has an important implication on the *cost effect* for asymmetric member countries (such as $\{ns^u\}$): while the benefit and cost of the north member country is larger relative to that of the south member country under $\{ns^u\}$ as under $\{ns\}$, the difference in costs of defection dominates the difference in benefits from defection under

$\{ns^u\}$ unlike $\{ns\}$. Thus, the north member country is less willing to cooperate relative to the south member country under $\{ns^u\}$: $\delta_n^{ns^u} > \delta_s^{ns^u}$. The nature of the *cost effect* for non-member countries under a bilateral CU remains the same as that under bilateral FTAs.

In the light of the above discussion it is immediate that the membership effect and the cost effect reinforce each other under $\{ss'^u\}$ and thus the south member countries' incentive is binding for the sustainability of multilateral cooperation under $\{ss'^u\}$: $\delta_s^{ss'^u} > \delta_n^{ss'^u}$ (represented by Figure 5a):

Proposition 9a *Under $\{ss'^u\}$, multilateral cooperation over free trade is sustainable if and only if $\delta > \delta_s^{ss'^u}$.*

[Insert Figure 5a]

Similarly, we show that the north member country's incentive determines the critical discount factor under $\{ns^u\}$ (see Figure 5b):

Proposition 9b *Under $\{ns^u\}$, multilateral cooperation over free trade is sustainable if and only if $\delta \geq \delta_n^{ns^u}$.*

[Insert Figure 5b]

Consider now $\{nn'^u\}$ where the membership effect and the cost effect work in opposite directions. We show that, when the south non-member countries are sufficiently high cost, the latter effect dominates the former and thus the south non-member countries' incentive binds for the multilateral cooperation over free trade: $\delta_s^{nn'^u} \geq \delta_n^{nn'^u}$ iff $c \geq c^{nn'^u}$. Therefore, the following is immediate (see Figure 5c):

Proposition 9c *Under $\{nn'^u\}$, multilateral cooperation over free trade is sustainable if and only if $\delta \geq \max(\delta_s^{nn'^u}, \delta_n^{nn'^u})$ where $\delta_s^{nn'^u} \geq \delta_n^{nn'^u}$ iff $c \geq c^{nn'^u}$.*

Insert Figure 5c

Next, using the above three propositions, we obtain one of our main results that compares the critical discount factors under different types of CUs.

Proposition 10 *The range of discount factors above which multilateral cooperation over free trade is sustainable is the largest under $\{nn'^u\}$ and the smallest under $\{ss'^u\}$: $\delta_s^{ss'^u} \geq \delta_n^{ns^u} \geq \max(\delta_s^{nn'^u}, \delta_n^{nn'^u})$, where equality obtains only when $c = 0$.*

The above proposition argues that the critical discount factors under different CUs follow a reverse ranking to that of aggregate world welfare. Specifically, $\{nn'^u\}$ not only yields the highest aggregate world welfare but also multilateral cooperation over free trade is more likely to be sustainable under $\{nn'^u\}$ relative to other CUs. Note that, when the degree of cost asymmetry across countries is not very large, this result provides a sharp contrast to Proposition 6. Under such a case, in terms of sustaining cooperation, while $\{nn'^u\}$ serves as the best agreement among bilateral CUs, $\{nn'\}$ is the worst agreement among bilateral FTAs.

4.5.1 Comparison of CUs with no Agreement

Does the formation of CUs hamper or facilitate multilateral cooperation relative to no agreement? To address this question, first note that the *membership effect* and the *cost effect* reinforce each other under $\{ss'^u\}$ so that multilateral cooperation is harder to sustain under $\{ss'^u\}$ relative to $\{\Phi\}$: $\delta_s^{ss'^u} > \delta_s^\Phi$ for all c . While the membership effect and cost effect work in opposite directions under $\{ns^u\}$, the former effect dominates the latter and the following obtains: $\delta_n^{ns^u} > \delta_s^\Phi$ for all c . Finally, consider $\{nn'^u\}$ which is the best possible CU in sustaining multilateral cooperation. We find that, when south countries are sufficiently high cost, the formation of $\{nn'^u\}$ facilitates multilateral cooperation relative to $\{\Phi\}$: $\delta_s^\Phi \geq \max(\delta_s^{nn'^u}, \delta_n^{nn'^u})$ iff $c \geq c(nn'^u - \Phi)$ holds. To understand this result, first note that $c(nn'^u - \Phi) < c^{nn'^u}$ holds. Therefore, we have to consider three distinct cost ranges: (i) when $c < c(nn'^u - \Phi)$ holds, $\max(\delta_s^{nn'^u}, \delta_n^{nn'^u}) = \delta_n^{nn'^u}$ and $\delta_n^{nn'^u} > \delta_s^\Phi$ obtain; (ii) when $c^{nn'^u} > c > c(nn'^u - \Phi)$ holds, $\max(\delta_s^{nn'^u}, \delta_n^{nn'^u}) = \delta_s^{nn'^u}$ and $\delta_s^{nn'^u} < \delta_s^\Phi$ obtain, and (iii) when $c > c^{nn'^u}$ holds, $\max(\delta_s^{nn'^u}, \delta_n^{nn'^u}) = \delta_s^{nn'^u}$ and $\delta_s^{nn'^u} < \delta_s^\Phi$ obtain. The intuition behind part (i) and part (ii) is as follows. When the degree of cost asymmetry is sufficiently small ($c < c(nn'^u - \Phi)$), the *cost effect* for south countries under $\{\Phi\}$ is small and dominated by the *membership effect* for the north member under $\{nn'^u\}$. The opposite obtains when c exceeds the critical threshold, i.e. when $c > c(nn'^u - \Phi)$ holds. Finally, part (iii) is an immediate implication of Proposition 4. We summarize these findings below.

Proposition 11 *While multilateral cooperation over free trade is always easier to sustain under $\{\Phi\}$ relative to $\{ns^u\}$ and $\{ss'^u\}$, it is harder to sustain under $\{\Phi\}$ relative to $\{nn'^u\}$ when $c \geq c(nn'^u - \Phi)$ holds.*

[Insert Figure 6]

Proposition 7 and Proposition 11 together imply that how the formation of PTAs affects multilateral cooperation over free trade depends on (i) whether a bilateral PTA is in the form of an FTA or a CU; (ii) the type of an FTA and a CU and (iii) the degree of cost asymmetry. Specifically, while a North-South PTA (FTA or CU) unambiguously hampers the obtainment of multilateral free trade, whether the formation of North-North and South-South PTAs facilitates the obtainment of free trade is ambiguous depending on the form of a PTA and the degree of cost asymmetry across regions.

Next, we consider a scenario in which we allow two bilateral PTAs.

5 Cooperation under two bilateral PTAs

Allowing for the formation of two bilateral PTAs yields two distinct trade agreement networks: (i) an inter-regional PTA network: two North-South PTAs (denoted by $\{ns, n's'\}$ if agreements are in the form of FTAs and $\{ns^u, n's'^u\}$ if they are CUs) and (ii) an intra-regional PTA network: North-North and South-South PTAs (denoted by $\{nn', ss'\}$ if agreements are in the form of FTAs and $\{nn'^u, ss'^u\}$ if they are CUs). First, we consider a static game and examine the following questions: Do non-member countries under an existing PTA have incentives to form an additional

PTA? What are the welfare implications of an additional PTA on the members of the other PTA and world as a whole?

Proposition 12 *Suppose there exist already a bilateral PTA. Then, the following results hold: (i) non-member countries under an existing PTA always have incentives to form a bilateral PTA with each other; (ii) while the formation of an additional PTA always raises the world welfare, it raises (lowers) the welfare of the member countries of the existing PTA if it is in the form of an FTA (a CU) and (iii) two inter-regional PTAs lead to higher world welfare relative to two intra-regional PTAs.*

Due to market segmentation, it is immediate that the intuition underlying the first two parts of the above proposition is the same as the one that underlies Proposition 2. The last part of the proposition argues in favor of the idea that, since trade creation is larger under inter-regional PTAs while trade diversion is smaller relative to intra-regional PTAs, the former regime welfare-dominates the polarization of the world trade regime with two intra-regional PTAs. Which of these regimes helps for the cause of multilateral cooperation over free trade? Hereafter, to conserve space, we directly state our main results and prove them in the appendix:

Proposition 13a *Suppose that PTAs are in the form of FTAs. Then, free trade can be sustained via multilateral cooperation (i) under $\{nn', ss'\}$ iff $\delta > \delta_s^{nn', ss'}$ holds; (ii) under $\{ns, n's'\}$ iff $\delta > \delta_s^{ns, n's'}$ holds. Moreover, multilateral cooperation is more likely to be sustainable under $\{ns, n's'\}$ relative to $\{nn', ss'\}$: $\delta_s^{nn', ss'} \geq \delta_s^{ns, n's'}$.*

[Insert Figure 7a]

Proposition 13b *Suppose that PTAs are in the form of CUs. Then, free trade can be sustained via multilateral cooperation (i) under $\{nn^u, ss^u\}$ iff $\delta > \delta_s^{nn^u, ss^u}$ holds; (ii) under $\{ns^u, n's^u\}$ iff $\delta > \delta_s^{ns^u, n's^u} = \delta_n^{ns^u, n's^u}$ holds. Moreover, multilateral cooperation is more likely to be sustainable under $\{ns^u, n's^u\}$ relative to $\{nn^u, ss^u\}$: $\delta_s^{nn^u, ss^u} \geq \delta_s^{ns^u, n's^u} = \delta_n^{ns^u, n's^u}$.*

[Insert Figure 7b]

The above two propositions are quite analogous with only one exception: while $\delta_s^{ns, n's'} \geq \delta_n^{ns, n's'}$ under FTAs, $\delta_s^{ns^u, n's^u} = \delta_n^{ns^u, n's^u}$ obtains under CUs. The joint determination of common external tariffs under CU yields that even though the benefits and costs are larger for the north member countries under $\{ns^u, n's^u\}$ relative to south member countries, the ratio of cost of defection and benefit from defection is the same and thus the same critical discount factors obtain: $\delta_s^{ns^u, n's^u} = \delta_n^{ns^u, n's^u} = \frac{2}{3}$. Unlike CUs, FTA members impose their external tariffs independently and the ratio of cost of defection to benefit from defection under $\{ns, n's'\}$ is smaller for the south member countries relative to north ones. Therefore, south member countries' incentives bind for the sustainability of cooperation under $\{ns, n's'\}$. More importantly, the above propositions along with Figures 7a and 7b imply that the polarization of the world trade system via intra-regional PTAs not only reduces world welfare but also makes free trade less likely to be sustainable via multilateral cooperation relative to inter-regional PTAs.

6 Conclusion

In this paper, we have considered how the formation of PTAs among cost-asymmetric North and South countries impacts non-member countries, world welfare and the sustainability of multilateral of free trade. By utilizing a four-country setup with multiple high cost and low cost countries, we were able to differentiate the roles of both cost (high versus low) and membership (member versus non-member) in bilateral FTAs and CUs, and for the prospects of global free trade. In our framework, the *cost effect* served to make high cost countries less willing to participate in multilateral free trade, independently of whether they were initially part of an FTA or a CU or a non-member. The feasibility of free trade then hinged upon the *membership effect*, as in some countries it worked in the same direction as the cost effect while in other countries it worked in the opposite direction.

In the static model, our model exhibits some common concerns regarding South-South agreements, particularly that these arrangements lead to higher levels of protection and lower welfare than North-North and North-South agreements. This result is more prominent when the initial PTA is a customs union than when it is a FTA. However, in terms of sustainability, we find certain cases where South-South agreements are more likely to lead to global free trade than North-North or North-South agreements. However, this is not always the case. For example, considering only CUs, South-South agreements are the least likely to result in free trade. In this way, we have shown that both the type of agreement (CU versus FTA) and the extent of cost variation between the North and South are critical in the evaluation of whether South-South trade agreements are beneficial or harmful for the prospects of global free trade.

7 Appendix

In this Appendix we provide the necessary supporting calculations and proofs.

7.1 Tariff and welfare levels

We begin by reporting welfare levels under different policy regimes as a function of tariffs and then derive optimum tariffs. Using the welfare and tariff levels reported below, we can easily obtain the formulae for optimum welfare levels under all possible regimes.

No agreement $\{\Phi\}$

$$w_n(\Phi) = \underbrace{\frac{(\frac{4\alpha-2c-3t_n^\phi}{5})^2}{2}}_{CS_n(\Phi)} + \underbrace{(\frac{\alpha+2c+3t_n^\phi}{5})^2}_{\pi_{nn}(\Phi)} + \underbrace{\sum_{z \neq n} (\frac{\alpha+2c-2t_z^\phi}{5})^2}_{\sum_{z \neq n} \pi_{nz}(\Phi)} + \underbrace{t_n^\phi (\frac{3\alpha-4c-6t_n^\phi}{5})}_{TR_n(\Phi)}$$

and

$$w_s(\Phi) = \underbrace{\frac{(\frac{4\alpha-2c-3t_s^\phi}{5})^2}{2}}_{CS_s(\Phi)} + \underbrace{(\frac{\alpha-3c+3t_s^\phi}{5})^2}_{\pi_{ss}(\Phi)} + \underbrace{\sum_{z \neq s} (\frac{\alpha-3c-2t_z^\phi}{5})^2}_{\sum_{z \neq s} \pi_{sz}(\Phi)} + \underbrace{t_s^\phi (\frac{3\alpha+c-6t_s^\phi}{5})}_{TR_s(\Phi)}$$

where

$$t_n^\Phi = t_{n'}^\Phi = \frac{9\alpha-2c}{33} \text{ and } t_s^\Phi = t_{s'}^\Phi = \frac{9\alpha-7c}{33}$$

Preferential Trade Agreements

North-North PTA - Suppose trade regime r represents a North-North FTA or CU: $r = \langle \{nn'\} \rangle$ if it is an FTA and $\langle \{nn'^u\} \rangle$ if it is a CU. Then, we have

$$w_n(r) = \underbrace{2(\frac{2\alpha-c-t_n^r}{5})^2}_{CS_n(r)} + \underbrace{\sum_{z=n,n'} (\frac{\alpha+2c+2t_z^r}{5})^2}_{\sum_{z=n,n'} \pi_{nz}(r)} + \underbrace{\sum_{z=s,s'} (\frac{\alpha+2c-2t_z^r}{5})^2}_{\sum_{z=s,s'} \pi_{nz}(r)} + \underbrace{2t_n^r (\frac{\alpha-3c-3t_n^r}{5})}_{TR_n(r)}$$

$$w_s(r) = \underbrace{\frac{(\frac{4\alpha-2c_s-3t_s^r}{5})^2}{2}}_{CS_s(r)} + \underbrace{(\frac{\alpha-3c_s+3t_s^r}{5})^2}_{\pi_{ss}(r)} + \underbrace{(\frac{\alpha-3c_s-2t_{s'}^r}{5})^2}_{\pi_{ss'}(r)} + \underbrace{\sum_{z=n,n'} (\frac{\alpha-3c_s-2t_z^r}{5})^2}_{\sum_{z=n,n'} \pi_{sz}(r)} + \underbrace{t_s^r (\frac{3\alpha+c_s-6t_s^r}{5})}_{TR_s(r)}$$

where $r = \langle \{nn'\} \rangle, \langle \{nn'^u\} \rangle$ and

$$\begin{aligned} t_n^{nn'} &= t_{n'}^{nn'} = \frac{\alpha-3c}{8} \text{ and } t_s^{nn'} = t_{s'}^{nn'} = \frac{9\alpha-7c}{33} \\ t_n^{nn'^u} &= t_{n'}^{nn'^u} = \frac{\alpha-c}{4} \text{ and } t_s^{nn'^u} = t_{s'}^{nn'^u} = \frac{9\alpha-7c}{33} \end{aligned}$$

South-South PTA - Suppose trade regime r represents a South-South FTA or CU: $r = \langle \{ss'\} \rangle$ if it is an FTA and $\langle \{ss'^u\} \rangle$ if it is a CU. Then, we have

$$w_n(r) = \underbrace{\frac{(\frac{4\alpha-2c-3t_n^r}{5})^2}{2}}_{CS_n(r)} + \underbrace{(\frac{\alpha+2c+3t_n^r}{5})^2}_{\pi_{nn}(r)} + \underbrace{(\frac{\alpha+2c-2t_{n'}^r}{5})^2}_{\pi_{nn'}(r)} + \underbrace{\sum_{z=s,s'} (\frac{\alpha+2c-3t_z^r}{5})^2}_{\sum_{z=s,s'} \pi_{nz}(r)} + \underbrace{t_n^r (\frac{3\alpha-4c-6t_n^r}{5})}_{TR_n(r)}$$

$$w_s(r) = \underbrace{2(\frac{2\alpha-c-t_s^r}{5})^2}_{CS_s(r)} + \underbrace{\sum_{z=s,s'} (\frac{\alpha-3c+2t_z^r}{5})^2}_{\sum_{z=s,s'} \pi_{sz}(r)} + \underbrace{\sum_{z=n,n'} (\frac{\alpha-3c-2t_z^r}{5})^2}_{\sum_{z=n,n'} \pi_{sz}(r)} + \underbrace{2t_s^r (\frac{\alpha+2c-3t_s^r}{5})}_{TR_s(r)}$$

where $r = \{ss'\}, \langle \{ss'^u\} \rangle$ and

$$\begin{aligned} t_s^{ss'} &= t_{s'}^{ss'} = \frac{\alpha+2c}{8} \text{ and } t_n^{ss'} = t_{n'}^{ss'} = \frac{9\alpha-2c}{33} \\ t_s^{ss'^u} &= t_{s'}^{ss'^u} = \frac{\alpha}{4} \text{ and } t_n^{ss'^u} = t_{n'}^{ss'^u} = \frac{9\alpha-2c}{33} \end{aligned}$$

North-South PTA - Suppose trade regime r represents a North-South FTA or CU: $r = \{ns\}$ if it is an FTA and $\{ns^u\}$ if it is a CU. Then, we have

$$\begin{aligned}
w_n(r) &= \underbrace{2\left(\frac{2\alpha - c - t_n^r}{5}\right)^2}_{CS_n(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha + 2c + 2t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{nz}(r)} + \underbrace{\sum_{z=n',s'} \left(\frac{\alpha + 2c - 2t_z^r}{5}\right)^2}_{\sum_{z=n',s'} \pi_{nz}(r)} + \underbrace{t_n^r \left(\frac{2\alpha - c - 6t_n^r}{5}\right)}_{TR_n(r)} \\
w_s(r) &= \underbrace{2\left(\frac{2\alpha - c - t_s^r}{5}\right)^2}_{CS_s(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha - 3c + 2t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{sz}(r)} + \underbrace{\sum_{z=n',s'} \left(\frac{\alpha - 3c - 2t_z^r}{5}\right)^2}_{\sum_{z=n',s'} \pi_{sz}(r)} + \underbrace{t_s^r \left(\frac{2\alpha - c - 6t_s^r}{5}\right)}_{TR_s(r)} \\
w_{n'}(r) &= \underbrace{\left(\frac{4\alpha - 2c - 3t_{n'}^r}{5}\right)^2}_2 + \underbrace{\left(\frac{\alpha + 2c + 3t_{n'}^r}{5}\right)^2}_{\pi_{n'n'}(r)} + \underbrace{\left(\frac{\alpha + 2c - 2t_{s'}^r}{5}\right)^2}_{\pi_{n's'}(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha + 2c - 3t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{n'z}(r)} + \underbrace{t_{n'}^r \left(\frac{3\alpha - 4c - 6t_{n'}^r}{5}\right)}_{TR_{n'}(r)} \\
w_{s'}(r) &= \underbrace{\left(\frac{4\alpha - 2c - 3t_{s'}^r}{5}\right)^2}_2 + \underbrace{\left(\frac{\alpha - 3c + 3t_{s'}^r}{5}\right)^2}_{\pi_{s's'}(r)} + \underbrace{\left(\frac{\alpha - 3c - 2t_{n'}^r}{5}\right)^2}_{\pi_{s'n'}(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha - 3c - 3t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{s'z}(r)} + \underbrace{t_{s'}^r \left(\frac{3\alpha + c - 6t_{s'}^r}{5}\right)}_{TR_{s'}(r)}
\end{aligned}$$

where $r = \{ns\}, \{ns^u\}$ and

$$\begin{aligned}
t_n^{ns} &= \frac{6\alpha + 7c}{48}; t_s^{ns} = \frac{6\alpha - 13c}{48} \text{ and } t_{n'}^{ns} = \frac{9\alpha - 2c}{33}; t_{s'}^{ns} = \frac{9\alpha - 7c}{33} \\
t_n^{ns^u} &= t_s^{ns^u} = \frac{2\alpha - c}{8} \text{ and } t_{n'}^{ns^u} = \frac{9\alpha - 2c}{33}; t_{s'}^{ns^u} = \frac{9\alpha - 7c}{33}
\end{aligned}$$

Inter-regional PTA network: $r = \{ns, n's'\}$ if it is an FTA and $\{ns^u, n's'^u\}$ if it is a CU. Then, we have

$$\begin{aligned}
w_n(r) &= \underbrace{2\left(\frac{2\alpha - c - t_n^r}{5}\right)^2}_{CS_n(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha + 2c + 2t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{nz}(r)} + \underbrace{\sum_{z=n',s'} \left(\frac{\alpha + 2c - 3t_z^r}{5}\right)^2}_{\sum_{z=n',s'} \pi_{nz}(r)} + \underbrace{t_n^r \left(\frac{2\alpha - c - 6t_n^r}{5}\right)}_{TR_n(r)} \\
w_s(r) &= \underbrace{2\left(\frac{2\alpha - c - t_s^r}{5}\right)^2}_{CS_s(r)} + \underbrace{\sum_{z=n,s} \left(\frac{\alpha - 3c + 2t_z^r}{5}\right)^2}_{\sum_{z=n,s} \pi_{sz}(r)} + \underbrace{\sum_{z=n',s'} \left(\frac{\alpha - 3c - 3t_z^r}{5}\right)^2}_{\sum_{z=n',s'} \pi_{sz}(r)} + \underbrace{t_s^r \left(\frac{2\alpha - c - 6t_s^r}{5}\right)}_{TR_s(r)}
\end{aligned}$$

where $r = \langle \{ns, n's'\} \rangle, \langle \{ns^u, n's'^u\} \rangle$ and

$$\begin{aligned}
t_n^{ns, n's'} &= t_{n'}^{ns, n's'} = \frac{6\alpha + 7c}{48} \text{ and } t_s^{ns, n's'} = t_{s'}^{ns, n's'} = \frac{6\alpha - 13c}{48} \\
t_n^{ns^u, n's'^u} &= t_{n'}^{ns^u, n's'^u} = t_s^{ns^u, n's'^u} = t_{s'}^{ns^u, n's'^u} = \frac{2\alpha - c}{8}
\end{aligned}$$

Intra-regional PTA network: $r = \langle \{nn', ss'\} \rangle$ if it is an FTA and $\langle \{nn'^u, ss'^u\} \rangle$ if it is a CU. Then, we have

$$w_n(r) = \underbrace{2\left(\frac{2\alpha - c - t_n^r}{5}\right)^2}_{CS_n(r)} + \underbrace{\sum_{z=n, n'} \left(\frac{\alpha + 2c + 2t_z^r}{5}\right)^2}_{\sum_{z=n, n'} \pi_{nz}(r)} + \underbrace{\sum_{z=s, s'} \left(\frac{\alpha + 2c - 3t_z^r}{5}\right)^2}_{\sum_{z=s, s'} \pi_{nz}(r)} + \underbrace{2t_n^r \left(\frac{\alpha - 3c - 3t_n^r}{5}\right)}_{TR_n(r)}$$

$$w_s(r) = \underbrace{2\left(\frac{2\alpha - c - t_s^r}{5}\right)^2}_{CS_s(r)} + \underbrace{\sum_{z=s, s'} \left(\frac{\alpha - 3c + 2t_z^r}{5}\right)^2}_{\sum_{z=s, s'} \pi_{sz}(r)} + \underbrace{\sum_{z=n, n'} \left(\frac{\alpha - 3c_s - 3t_z^r}{5}\right)^2}_{\sum_{z=n, n'} \pi_{sz}(r)} + \underbrace{2t_s^r \left(\frac{\alpha + 2c - 3t_s^r}{5}\right)}_{TR_s(r)}$$

where $r = \{nn', ss'\}$, $\{nn'^u, ss'^u\}$ and

$$\begin{aligned} t_n^{nn', ss'} &= t_{n'}^{nn', ss'} = \frac{\alpha - 3c}{8} \text{ and } t_s^{nn', ss'} = t_{s'}^{nn', ss'} = \frac{\alpha + 2c}{8} \\ t_n^{nn'^u, ss'^u} &= t_{n'}^{nn'^u, ss'^u} = \frac{\alpha - c}{4} \text{ and } t_s^{nn'^u, ss'^u} = t_{s'}^{nn'^u, ss'^u} = \frac{\alpha}{4} \end{aligned}$$

The proofs of Proposition 1, Proposition 2, Proposition 12 and the inequalities in (16) and (17) are immediate using the optimum tariffs and welfare levels reported above.

Critical Cost Thresholds

No Agreement

$$x_{sn}(\Phi) = \frac{3\alpha - 19c}{33} \geq 0 \text{ iff } c \leq \frac{3\alpha}{19}$$

Free Trade Agreements

$$x_{sn}(ss') = x_{sn'}(ss') = x_{sn}(n's') = \frac{3\alpha - 19c}{33} \geq 0 \text{ iff } c \leq \frac{3\alpha}{19}$$

$$x_{sn}(nn') = x_{sn}(nn', ss') = \frac{\alpha - 3c}{8} \geq 0 \text{ iff } c \leq \frac{\alpha}{3}$$

$$x_{sn'}(n's') = x_{sn'}(ns, n's') = \frac{2\alpha - 11c}{16} \geq 0 \text{ iff } c \leq \frac{2\alpha}{11}$$

$$x_{ss'}(n's') = x_{ss'}(ns, n's') = \frac{2\alpha - 7c}{16} \geq 0 \text{ iff } c \leq \frac{2\alpha}{7}$$

Based on these threshold cost levels, the market access of south firms under any FTA is guaranteed if $c \leq \bar{c}^f = \frac{3\alpha}{19}$ holds.

Customs Unions

$$x_{sn}(ss'^u) = x_{sn'}(ss'^u) = x_{sn}(n's'^u) = \frac{3\alpha - 19c}{33} \geq 0 \text{ iff } c \leq \frac{3\alpha}{19}$$

$$x_{sn}(nn'^u) = x_{sn}(nn'^u, ss'^u) = \frac{\alpha - 9c}{20} \geq 0 \text{ iff } c \leq \frac{\alpha}{9}$$

$$x_{sn'}(n's'^u) = x_{ss'}(n's') = x_{sn'}(ns^u, n's'^u) = x_{ss'}(ns^u, n's'^u) = \frac{2\alpha - 21c}{40} \geq 0 \text{ iff } c \leq \frac{2\alpha}{21}$$

Based on these threshold cost levels, the market access of south firms under any CU is guaranteed if $c \leq \bar{c}^u = \frac{2\alpha}{21}$ holds. Note that $\bar{c}^f < \bar{c}^u$. In (17), $c_s^{ns} \cong 0.086\alpha < \bar{c}^u$.

8 References

References

- [1] Aghion, Philippe, Pol Antràs, and Elhanan Helpman (2007). “Negotiating free trade.” *Journal of International Economics* 73, 1–30.

- [2] Bagwell, Kyle, and Robert W. Staiger (1997a). "Multilateral Cooperation during the Formation of Customs Unions," *Journal of International Economics* 42 (1997a), 91–123.
- [3] Bagwell, Kyle, and Robert W. Staiger (1997b). "Multilateral tariff cooperation during the formation of free trade areas." *International Economic Review* 38, 291–319.
- [4] Bagwell, Kyle, and Robert W. Staiger (1998). "Will Preferential Agreements Undermine the Multilateral Trading System?" *Economic Journal* 108, 1162–82.
- [5] Bagwell, Kyle, and Robert W. Staiger. "Regionalism and multilateral tariff cooperation." In: John Pigott and Alan Woodland, eds. *International Trade Policy and the Pacific Rim* (Macmillan, London), 1999.
- [6] Bhagwati, Jagdish. *The World Trading System at Risk* (Princeton University Press, Princeton, NJ.), 1991.
- [7] Bhagwati, J. and Arvind Panagariya , 1996, *Preferential Trade Areas and Multilateralism: Strangers, Friends or Foes?*, Bhagwati, Jagdish and Panagariya, Arvind (eds), *The Economics of Preferential Trade Agreements*, AEI Press, Washington DC, 83-100.
- [8] Bond, Eric W., Constantinos Syropoulos, and Alan L. Winters (2001). "Deepening of regional integration and multilateral trade agreements." *Journal of International Economics* 53, 335–362.
- [9] Brander, James A., and Paul R. Krugman (1983). "A 'Reciprocal Dumping' Model of International Trade," *Journal of International Economics* 15, 313–23.
- [10] Chang, Won and L. Alan Winters (2002). "How Regional Blocs Affect Excluded Countries: The Price Effects of MERCOSUR," *American Economic Review* 92, 889-904.
- [11] Das, S. and Subhadip Ghosh (2006), "Endogenous Trading Bloc Formation in a North-South Global Economy" *Canadian Journal of Economics*, 39 (3), 809-830.
- [12] Demir, Firat and Omar S. Dahi (2011). "Asymmetric effects of financial development on South-South and South-North trade: Panel data evidence from emerging markets," *Journal of Development Economics* 94, 139-149.
- [13] Estevadeordal, Antoni, Caroline Freund, and Emanuel Ornelas (2008). "Does regionalism affect trade liberalization toward non-members?" *Quarterly Journal of Economics* 123, 1531–1575.
- [14] Furusawa, Taiji, and Hideo Konishi (2007). "Free trade networks." *Journal of International Economics* 72, 310–335.
- [15] Goyal, Sanjeev, and Sumit Joshi (2006). "Bilateralism and free trade." *International Economic Review* 47, 749–778.

- [16] Karacaovali, Baybars and Nuno Limão (2008). “The clash of liberalizations: Preferential vs. multilateral trade liberalization in the European Union,” *Journal of International Economics* 74(2), 299-327.
- [17] Krishna, Pravin and Devashish Mitra (2008). “Reciprocated unilateralism in trade reforms with majority voting,” *Journal of Development Economics* 85, 81-93.
- [18] Limão, Nuno (2006). “Preferential Trade Agreements as Stumbling Blocks for Multilateral Trade Liberalization: Evidence for the United States,” *American Economic Review* 96(3), 896-914.
- [19] Magee, Stephen P., and Hak-Loh Lee (2001). “Endogenous Tariff Creation and Tariff Diversion in a Customs Union,” *European Economic Review* 45(3), 495–518.
- [20] Riezman, Raymond (1991). “Dynamic Tariffs with Asymmetric Information,” *Journal of International Economics* 30, 267–83.
- [21] Rodrik, D., 1994. The rush to free trade. Why so late? Why now? Will it last? In: Haggard, S., Webb, S.B. (Eds.), *Voting For Reform: Democracy, Political Liberalization, and Economic Adjustment* (Oxford University Press, New York).
- [22] Saggi, Kamal, (2006). “Preferential trade agreements and multilateral tariff cooperation,” *International Economic Review* 47, 29–57.
- [23] Saggi, Kamal, and Halis Murat Yildiz (2010). “Bilateralism, multilateralism, and the quest for global free trade.” *Journal of International Economics* 81, 26-37.
- [24] Saggi, Kamal, Alan Woodland, Halis Murat Yildiz (2013). “On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions,” *American Economic Journal: Microeconomics*, 5(1), 63-99.
- [25] Winters, L. Alan (1997). “Regionalism and the Rest of the World: The Irrelevance of the Kemp-Wan Theorem,” *Oxford Economics Papers*, 49, 228-34.
- [26] World Bank (2005), *Global Economic Prospects, Trade, Regionalism, and Development* (Washington, DC: World Bank).
- [27] World Trade Organization (2011). *World Trade Report 2011*.