Preferential versus Multilateral Trade Liberalization and the Role of Political Economy

Andrey Stoyanov[*]and Halis Murat Yildiz[†]

Abstract

In this paper we analyze the effect of the freedom to pursue preferential trade liberalization, permitted by Article XXIV of the GATT, on country's incentives to participate in multilateral negotiations and on the feasibility of the global free trade. We present a model in which countries choose whether to participate in preferential or multilateral trade agreements under political pressures from domestic special interest groups. We show that heterogeneity in political preferences across countries plays an important role for the relative merits of preferential and multilateral approaches to trade liberalization. On one hand, the opportunity to liberalize preferentially may be necessary to induce countries with strong political motivations to participate in multilateral free trade negotiations. On the other hand, when countries share similar political preferences, multilateral free trade that would have been politically supported otherwise becomes unattainable if countries can pursue preferential liberalization.

[*]Department of Economics, York University, 1084 Vari Hall, 4700 Keele St, Toronto, M3J 1P3, Canada. Tel: +1-416-736-2100 (ext. 22833); E-mail: andreyst@yorku.ca.

[†]Department of Economics, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3. Phone: 1-416-979-5000 (ext 6689); e-mail: hyildiz@ryerson.ca.

1   Introduction

Since the establishment of the General Agreement on Tariffs and Trade (GATT), trade reforms have proceeded along three main fronts: unilaterally, preferentially with a few partners, and multilaterally within the GATT, now subsumed by the World Trade Organization (WTO). Of the three approaches, preferential trade liberalization gained the most popularity in recent years, with the number of preferential trade agreements (PTAs) increased four-fold since completion of the Uruguay round of WTO negotiations in 1994.[1] In contrast, the last round of the multilateral trade talks – the Doha Round – have failed to come to a fruitful conclusion despite thirteen years of intense negotiations. Although PTAs are sanctioned by the Article XXIV of the GATT in order to promote freer trade, they come into direct conflict with the main GATT/WTO principles of reciprocity and non-discrimination. As such, there are concerns that the spread of PTAs may hurt the multilateral trading system and serve as an alternative, rather than a complement, to multilateral trade liberalization.[2]

Diverse interests of the WTO members are often blamed for the decline in multilateralism and proliferation of PTAs. In particular, differences in political idiosyncrasies and protectionist preferences across countries may induce "similar" countries to negotiate PTAs rather than participate in large scale multilateral deals. Therefore, the political economy factors may play an important role in the analysis of PTAs, and our goal in this paper is to study how heterogeneity in political preferences across countries affects the global trade system. We devise a model of endogenous formation of trade agreements – both bilateral and multilateral – with politically motivated governments in order to analyze whether or not the freedom to pursue Free Trade Agreements (FTAs), permitted by Article XXIV, is *necessary* to attain global free trade.[3] We find that while the ability to form FTAs reduces the likelihood of global free trade when countries have symmetric political preferences, with asymmetric preferences the global free trade may become politically feasible only when countries have the freedom to pursue FTAs. These results suggest that, in order to assess the effect of Article XXIV on the prospect of global free trade, political asymmetries across countries must be taken into account in the analysis of trade agreement formation.

The debate about the effect of PTAs on the trade system centers around the terms

---

[1] As of 15 June 2014, 585 notifications of PTAs (counting goods and services separately) had been received by the World Trade Organization (WTO). Of these notifications, 379 PTAs are already in force with others scheduled for implementation in the near future.

[2] While the most favored nation (MFN) clause requires the WTO member countries to undertake trade liberalization on a non-discriminatory basis, Article XXIV of the GATT permits a group of countries to pursue PTAs under which they can grant tariff (and other trade policy) concessions to each other that they do *not* have to extend to other WTO members.

[3] We do not consider unilateral trade liberalization since all countries have unilateral incentive to extract rent and such liberalization is not in any country's interest.

building and *stumbling blocs,* introduced by Bhagwati (1991). The standard view in the existing literature is that PTAs are building (stumbling) blocks to global free trade if the process of bilateral trade liberalization eventually converges (fails to converge) to global free trade.[4] However, this view can be misleading if global free trade obtains regardless of whether the formation of PTAs is allowed or not. In this paper, we take an alternative view that PTAs are building blocks to global free trade when bilateral trade agreements lay the foundation for multilateral free trade in the sense that the freedom to pursue PTAs is necessary in order for global free trade to be achieved.[5] Our approach to trade agreement formation follows that of Saggi and Yildiz (2010) under which both preferential and MFN-based multilateral agreements emerge endogenously. Using this approach in a political economy context, we take the building vs. stumbling blocs debate one step further. Specifically, we show that under certain conditions, the freedom to pursue FTAs (Article XXIV) is *necessary* to attain global free trade. In such a case, FTAs act as a *strong building bloc*. We also find that there are conditions under which the freedom to pursue FTAs *prevents* the attainment of global free trade that would arise when countries were to follow trade liberalization on an MFN basis only. When that happens, FTAs act as a *strong stumbling bloc*. To the best of our knowledge, this paper is the first attempt to address the role of "politics" in a model where the formation of trade agreements (both preferential and multilateral) is endogenous.

Our theoretical framework is based on a three-country model of trade with oligopolistic markets where countries optimally choose tariff rates and membership in trade agreements.[6] As in Grossman and Helpman (1995), policymakers in all countries are politically biased and balance social welfare with industrial interests. Our analysis draws on a comparison of equilibrium outcomes under bilateralism, where countries have a freedom to form FTAs, and multilateralism, where FTAs are not available. To this end, we employ a three stage game under both bilateralism and multilateralism. In the first stage of each game, equilibrium trade agreements are formed between countries. In the second stage countries choose import tariffs under a given trade policy regime, and in the

---

[4] Earlier literature on the building versus stumbling bloc discussion has taken trade agreements exogenously and focused on the incentives of countries to participate in multilateral free trade (see Grossman and Helpman, 1995; Krishna, 1998; Ornelas, 2005a,b). More recent studies, including Goyal and Joshi (2006), Aghion et al. (2007), Furusawa and Konishi (2007), and Seidman (2009) use endogenous formation of PTAs but ignore the possibility of trade liberalization on an MFN basis, and view PTAs as building blocs so long as their pursuit eventually leads to global free trade.

[5] Similarly, we argue that PTAs are stumbling blocs if the freedom to pursue bilateral FTAs hampers multilateral free trade that would have been obtained if countries were to pursue trade liberalization only multilaterally.

[6] Using the same market structure, Krishna (1998) demonstrates that, for exogenously given tariffs, formation of an FTA between two countries reduces their incentives to liberalize trade with respect to the third country. With endogenous tariffs, Ornelas (2005a,b) shows that an FTA induces member countries to lower their tariffs on the non-member country. In contrast to our paper, Krishna (1998) and Ornelas (2005a,b) do not develop an equilibrium theory of FTAs and only consider the effects of an exogenously given FTA.

third stage firms compete in a Cournot fashion. For the equilibrium analysis, we utilize the concept of coalition-proof Nash equilibrium (CPNE) and focus on Nash agreements that are immune to self-enforcing coalitional deviations. This refinement of Nash equilibrium is important in order to capture the negotiations of trade agreements in a more realistic fashion. Consistent with the definition of CPNE in Bernheim et al. (1987), all countries are free to arrange mutually beneficial deviations from Nash agreements and an initial coalitional deviation from a Nash equilibrium must be self-enforcing, i.e. taking the strategy of the complement as fixed, no proper subset of the deviating coalition has an incentive for further deviation.[7] By comparing the sets of CPNE agreements under bilateralism and multilateralism, we identify the consequences of the exception to non-discrimination principle incorporated in the GATT Article XXIV.

The model generates new insights into the optimal tariff formation under both preferential and two-country multilateral agreements (MA) with political biases. As a starting point, our optimal tariff discussion under a bilateral FTA confirms the findings by Ornelas (2005a,b) about the effect of political economy forces on trade policy: while the optimal tariffs rise with the government's political bias under all trade regimes, an FTA induces each member to reduce external tariffs against non-members and the reduction is deeper for more politically motivated governments. We complement Ornelas' results by showing that the same result obtains under a MA. This finding captures the free-riding problem inherent to multilateral trade liberalization which does not involve all countries since the non-participating country benefits from the multilateral trade liberalization undertaken by the other two countries without having to offer any liberalization in return. More importantly, the comparison of external tariffs of member countries under an FTA and an MA reveals that the degree of trade liberalization undertaken by member countries is lower when they sign a multilateral agreement relative to the one when they sign a preferential agreement. This gap between external tariffs widens as the member countries become more politically motivated. This gap determines the payoff of the non-member country under these two alternative trade agreements and hence the country's incentives to participate in one or the other.

The relative effect of an FTA and an MA between a pair of countries on a third

---

[7]In examining endogenous formation of PTAs, Goyal and Joshi (2006) and Furusawa and Konishi (2007) employ the concept of pairwise stability in a network formation game developed by Jackson and Wolinsky (1996). Unlike CPNE, the concept of pairwise stability implies two constraints. First, the deviating coalition can contain at most two countries. Second, the deviation can consist of severing just one existing link or forming one additional link. Unlike the present paper, they only examine whether the formation of FTAs results in a global free trade as the stable outcome and do not analyze the consequences of adopting a multilateral approach to trade liberalization. Like us, Riezman (1999) also asks whether bilateralism facilitates or hinders the achievement of global free trade. However, while we analytically derive CPNE agreements, Riezman (1999) uses the cooperative solution concept of the core and illustrates his results via numerical examples. Saggi and Yildiz (2010) and Saggi et al. (2013) use the CPNE equilibrium concept but do not analyze the role of political economy factors in trade agreement formaion.

country is determined by two opposing forces. First, there is a *tariff level effect* which makes the outside country better off under the FTA because preferential agreements result in lower tariffs for non-members relative to MA. Second, there is a *discrimination effect* that exists only under a bilateral FTA and lowers the outside country's payoff. The discrimination effect stems from the fact that the FTA tariffs only apply to a non-member country which thus is subject to a discriminatory treatment in each member country's market under a bilateral FTA. The combination of these two effects yields that the payoff of the non-member country's government is lower under an FTA than under an MA when the political biases of the member countries are sufficiently low. Intuitively, as the political biases of member countries rise, the tariff complementarity deepens and the tariff level effect becomes stronger, while the discrimination effect weakens. This fundamental difference between bilateral and multilateral trade liberalization plays a crucial role in our analysis of equilibrium trade agreements that we examine next.

In analyzing equilibrium trade agreements, we start with symmetric political preferences and find that the equilibrium outcome is determined by the strength of the political bias. First, when governments have sufficiently strong political biases, free trade fails to be a CPNE under both bilateralism and multilateralism. This finding contrasts with most of the previous literature on endogenous trade agreements without political economy factors. For example, in a model with symmetric countries and welfare-maximizing governments, Saggi and Yildiz (2010) argue that free trade obtains as a unique CPNE regardless of the trade regime in place. Second, when the degree of political motivation is sufficiently low, free trade is a unique CPNE under both trade regimes. This result converges to the main finding by Goyal and Joshi (2006), who analyze endogenous formation of trade agreements in the oligopoly model of trade where countries pursue PTAs only and governments have no political biases. While they interpret this result as compatibility of bilateralism and free trade, we show that it may break down if governments have political motivations. Third, when the political biases are at the intermediate range, the ability to form FTAs reduces the likelihood of global free trade arising as a CPNE. As discussed above, when such is the case, FTAs act as a strong stumbling bloc for multilateral free trade. Over this range of political bias parameter, the tariff level effect dominates the discrimination effect and each country has an incentive to free ride on the bilateral trade liberalization by the other two countries, while such an incentive does not exist under multilateralism. Finally, when global free trade is out of reach under both bilateralism and multilateralism, we find that the ability to form bilateral FTAs can either improve (*partial building bloc*) or reduce (*partial stumbling bloc*) global welfare, depending on the degree of political motivation.[8]

---

[8]Partial building bloc result stems from the fact that trade liberalization is deeper under an FTA relative to MA and thus the former yields higher global welfare. When political biases are sufficiently strong, 'no agreement' arises under bilateralism while an MA obtains under multilateralism. When such

Next, we examine the scenario where countries have asymmetric political preferences. Empirical tests of the Grossman and Helpman (1994)'s protection for sale model consistently report substantial heterogeneity in political preferences across countries.[9] In a recent study, Gawande et al. (2009) estimate the weight that governments attach to social welfare relative to industry lobbying for a large number of countries in a unified theoretical and empirical framework. They report large variation in political biases across countries, with the third quartile of the political bias parameter distribution being seven-fold greater than the first quartile. Furthermore, Krugman (1991) and Grossman and Helpman (1995) noted that asymmetries across countries can play a crucial role in determining incentives for preferential and multilateral trade liberalization. Yet the existing literature on trade agreements has tended to pay little attention to cross-country heterogeneity in political preferences.[10] We introduce this heterogeneity in our model by assuming that one country has stronger political bias than the other two (symmetric) countries. Our results show that heterogeneity across countries may be an important determinant of the success of multilateralism and that bilateralism may facilitate the process of global trade liberalization.

Under asymmetric political preferences, both the degree of political biases and the heterogeneity in political preferences determine the equilibrium outcomes under both bilateralism and multilateralism. Consistent with this, whether global free trade is a CPNE or not depends critically on the unilateral deviation incentive of the country with the strongest political bias. When governments' political biases are sufficiently weak and the degree of asymmetry in political preferences are sufficiently small, free trade obtains under both games and thus the freedom to pursue bilateral FTAs is irrelevant for the ultimate objective of achieving global free trade. More interestingly, we show that unlike the symmetric case, the freedom to pursue FTAs can be necessary for achieving global free trade when the asymmetry in political preferences is sufficiently large and the political biases of countries are not very strong. Under such a case, the country with the strongest political preference has incentive to unilaterally deviate from free trade under multilateralism but not under bilateralism since the discrimination effect dominates the tariff level effect. Thus, the threat of a bilateral FTA between two countries and the discrimination that is inherent to such trade agreement can be necessary to convince the country with the strongest political preference to announce in favor of the global free trade. When such is the case, Article XXIV acts as a strong building bloc for

is the case, we argue that partial trade liberalization is better than none and thus partial stumbling blocs result obtains.

[9] See Gawande and Bandyopadhyay (2000) for the U.S., Mitra et al. (2002) for Turkey, Cadot et al. (2003) for India, McCalman (2004) for Australia, Stoyanov (2009) for Canada.

[10] Furusawa and Konishi (2007) consider asymmetry in the size of the industrial good industry and the market size. In Saggi and Yildiz (2010) the underlying asymmetry is in the endowment levels. Political economy factors play no role in these papers.

achieving the global free trade. However, when governments are relatively symmetric in their political biases which are moderately strong (or as we converge to the complete symmetry), the tariff level effect dominates the discrimination effect and the incentives for participating in global free trade are stronger when discriminatory bilateral agreements are not permitted. In this case, bilateral FTAs act as a strong stumbling bloc for the global free trade. Finally, when global free trade is out of reach under both bilateralism and multilateralism and bilateral FTAs and MAs are CPNE, the option to pursue bilateral FTAs can yield a welfare-improving trade liberalization that is foregone under a strictly multilateral approach and thus bilateral FTAs act as a partial building bloc.[11]

The paper proceeds as follows. Section 2 presents the underlying trade model with political economy discussion. Sections 3 and 4 present the equilibrium analysis under bilateralism and multilateralism with symmetric political preferences. Section 5 extends the equilibrium analyses under these two regimes to the case with asymmetric political preferences across countries and then Section 6 concludes. Proofs are collected in Appendix A. The implications of alternative pattern of asymmetry are discussed in Appendix B.

2   Model

We develop a simple oligopoly model of trade in which each country has a unilateral incentive to impose a rent-extracting tariff on those trade partners with whom it does not have any trade agreement. There are three countries in the model, $i, j, k \equiv a, b, c$, two sectors, producing goods y and x, and a single factor of production (labor, L). Good y is a numéraire produced in every country by perfectly competitive firms under constant return to scale production technology with labor productivity normalized to one. The numéraire good is freely traded so that wages are equal to unity in all countries and total income is equal to the labor supply.[12] Good x is produced in each country by a single profit-maximizing firm at a constant marginal cost in terms of the numeraire good y. Production technologies are identical across countries. Therefore, international trade results only from the oligopolistic behavior in sector x.[13]

A representative consumer in each country maximizes a quasilinear utility given by

$$U(x_i, y_i) = u(x_i) + y_i \tag{1}$$

---

[11] As the degree of asymmetry falls and political biases are strong, FTAs can become a partial stumbling bloc, converging to the symmetric case.

[12] The numéraire good is freely traded across countries to settle the balance of trade.

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

where $x_i \equiv \sum\limits_{z=i,j,k} x_{zi}$ is the total consumption of good x in country i and x_{zi} is country i's consumption of x produced in country z. We also assume that the non-linear part of the utility function is quadratic in x:

$$u(x_i) = \alpha x_i - \frac{x_i^2}{2}$$

It follows that the inverse demand function in country i is given by:

$$p_i(x_i) = \alpha - x_i \tag{2}$$

where α represents the market size of each country.[14]

In the absence of any trade agreement, firm j faces a specific tariff t_{ji} when exporting to country i, with $t_{ii} = 0$ for all i. Denote the vector of tariff schedules of country i by $\mathbf{t}_i \equiv (t_{ij}, t_{ik})$. Since tariff discrimination is prohibited by Article I of the GATT, we restrict attention to the case where $t_{ij} = t_{ik} = t_i$ for all i, j, k when there is no FTA in place. These trade barriers provide a cost advantage to domestic firms in local markets by increasing the effective unit cost serving the market for foreign firms:

$$c_{zi} = c + t_{iz} \tag{3}$$

Then the profit function of firm z from sales in country i can be written as:

$$\pi_{zi} = x_{zi} p_i(x_i) - c_{zi} x_{zi} \tag{4}$$

The first-order conditions for profit maximization yield equilibrium output levels:

$$x_{ii} = \frac{\alpha - 3c + \sum\limits_{z \neq i} c_{zi}}{4} \text{ and } x_{ji} = \frac{\alpha + c - 3c_{ji} + c_{ki}}{4}, \ z \neq i \tag{5}$$

Equilibrium price in country i is easily calculated from the demand function:

$$p_i = \frac{\alpha + c_i + \sum\limits_{z \neq i} c_{zi}}{4} \tag{6}$$

Using (5) and (6), equilibrium profits are:

$$\pi_{ii} = x_{ii}^2 \text{ and } \pi_{zi} = x_{zi}^2, \ z \neq i \tag{7}$$

Import tariffs protect the import-competing firms and raise their domestic profits at

---

[14] We abstract from asymmetries in country sizes and technologies in order to focus on the implications of heterogeneity in political preferences across countries for the equilibrium trade policy.

expense of foreign firms, so the following comparative statics results apply:

$$\frac{\partial \pi_{ji}}{\partial t_{ji}} = -\frac{3x_{ji}}{2} < 0, \frac{\partial \pi_{ii}}{\partial t_{ij}} = \frac{x_{ii}}{2} > 0, \text{ and } \frac{\partial \pi_{ki}}{\partial t_{ji}} = \frac{x_{ki}}{2} > 0. \tag{8}$$

Let country i's national welfare function, denoted by W_i, consist of consumer surplus, producer surplus, and tariff revenue:

$$W_i = u(x_i) - x_i p_i(x_i) + \pi_i + \sum_{z \neq i} t_{iz} x_{zi} \tag{9}$$

where producer surplus (π_i) is the sum of domestic firm's profits in the local market (denoted by π_{ii}) and the foreign markets, denoted by $\sum_{z \neq i} \pi_{iz}$: $\pi_i = \pi_{ii} + \sum_{z \neq i} \pi_{iz}$. It is useful to define domestic surplus (denoted by DS_i) as the difference between the national welfare and export profits: $DS_i = W_i - \sum_{z \neq i} \pi_{iz}$. Note that due to market segmentation, strategic independence of trade policies obtains and thus DS_i is fully determined by the own tariff $\mathbf{t}_i$, while export profits depend on foreign tariffs only.

Following Grossman and Helpman (1994), we define government's preferences as the weighted sum of national welfare and political contributions received from domestic lobby groups (C_i): each government values a dollar received as contributions potentially more than a dollar added to national welfare. Let β_i to be the degree of a government's political bias, measured by the extent of government's preference for contributions vis-a-vis national welfare. Thus, the government's payoff function can be expressed as:

$$g_i(\mathbf{t}_i, C_i) = W_i(\mathbf{t}_i) + \beta_i C_i$$

We assume that each oligopolistic firm is politically active and can provide its own government with political contributions. Given contributions C_i, each firms in the oligopolistic market maximizes its net payoff given by the difference between its total profits across all markets and the contributions it provides to the local government:[15]

$$\psi_i(\mathbf{t}_i, C_i) = \pi_i(\mathbf{t}_i) - C_i$$

We follow Maggi and Rodriguez-Clare (1998) and Ornelas (2005a,b) and model the lobbying process as a simple Nash-bargaining game between the local firm and the government.[16] The equilibrium in this bargaining process obtains for the tariff that

---

[15]Similar to Grossman and Helpman (1995), we assume that lobbying takes place in each country independently. Firm in each country lobby the national government and it is not possible to lobby internationally. For detailed discussion of the bargaining process and equilibrium conditions, see Ornelas (2005b,c).

[16]Baldwin (1987) argues that this formulation is consistent with the reduced form of many distinct ways (including the "protection for sale" model by Grossman and Helpman, 1994) to incorporate general political and distributional concerns in the governments' objectives.

maximizes the sum of government's and firm's objective functions, or equivalently for the tariff which solves the following problem:

$$\max_{\mathbf{t}_i} g_i(\mathbf{t}_i) \equiv DS_i + \beta_i \pi_{ii} + \sum_{z=j,k} \pi_{iz}$$

Under no trade agreement, governments choose tariffs in anticipation of the firm's contributions, which in turn depend on the chosen tariff levels. When a bilateral FTA is formed, tariffs between member countries are set to zero and the firm bargains with the local government over the external tariff on the non-member country. Under a multilateral free trade all tariffs are eliminated and there is no further policy to be enacted. Thus, lobbying does not take place ($\beta_i = 0$), leading to the equivalence of the government's payoff and national welfare.

3 Endogenous agreements under bilateralism

We now describe our three stage game of trade liberalization, which we refer to as *bilateralism*, when governments have an option to form bilateral FTAs. In the first stage, each country simultaneously announces whether or not it wants to sign an FTA with each of its trading partners. It is important to emphasize that in this game countries do not announce in favor of a specific trade agreement but rather name partners with whom they wants to form trade agreements. Country i's announcement regarding its preferred trade regime with countries j and k is denoted by σ_i and its strategy set Ω_i consists of four possible announcements:

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

where the announcement $\{\phi, \phi\}$ is in favor of the status quo (no FTA with either trading partner), $\{j, \phi\}$ in favor of an FTA with only country j; $\{\phi, k\}$ in favor of an FTA only with country k; and $\{j, k\}$ in favor of FTAs with both trading partners. The above first stage determines the underlying trade policy regime. Given the policy regime, governments impose their optimal external tariffs in the second stage. Finally, for a given policy regime and import tariffs, firms compete in the product markets in a Cournot fashion. Since markets are segmented, firms make independent decisions regarding their sales in each market (as in Brander and Krugman, 1983).

The following policy regimes can emerge in the bilateralism game: (i) no agreement or the status quo, $\langle \Phi \rangle$, would prevail if no two announcements match or when everyone announces $\{\phi, \phi\}$; (ii) an FTA between countries i and j, denoted by $\langle ij \rangle$, would form if and only if i and j announce in favor of FTA with each other and there is no other matching announcement: i.e., $i \in \sigma_j$ and $j \in \sigma_i$ while $i \notin \sigma_k$ and/or $k \notin \sigma_i$ and $j \notin \sigma_k$

and/or $k \notin \sigma_j$; (iii) two independent FTAs (i.e. a hub and spoke trade regime) in which i is the common member, denoted by $\langle ih \rangle$, are formed if and only if country i announces in favor of signing an FTA with countries j and k while countries j and k announce only in favor of FTA with country i: i.e. $j \in \sigma_i$ and $i \in \sigma_j$ and $k \in \sigma_i$ and $i \in \sigma_k$ while $k \notin \sigma_j$ and/or $j \notin \sigma_k$ and (iv) free trade, denoted by $\langle F \rangle$, would obtain if and only if all countries announce each others' names.[17]

In what follows, we first derive the set of Nash equilibria and then allow countries to undertake coalitional deviations in order to isolate the set of coalition-proof Nash equilibria (CPNE). Bernheim et. al. (1987) state that "in an important class of "noncooperative" environments, it is natural to assume that players can freely discuss their strategies, but cannot make binding commitments. In such cases, any meaningful agreement between the players must be self-enforcing. Although the Nash best-response property is a necessary condition for self-enforceability, it is not sufficient – it is in general possible for coalitions to arrange plausible, mutually beneficial deviations from Nash agreements." The application of the CPNE solution concept is eminently desirable in the present context since countries considering bilateral trade agreements certainly have the capacity to communicate with one another without necessarily having the ability to make binding commitments regarding their future plans with respect to such trade agreements.

Before deriving equilibrium trade agreements, we clarify an expositional point: while changes in the underlying trade regime result from announcement deviations by governments, it proves more convenient to refer directly to regime changes rather than changes in announcements. For example, when the bilateral FTA $\langle ij \rangle$ is in place, the unilateral announcement deviation of government i from $\{j, \phi\}$ to $\{\phi, \phi\}$ alters the underlying trade regime from $\langle ij \rangle$ to no agreement $\langle \Phi \rangle$, and we refer to this announcement deviation of government i as simply a deviation from $\langle ij \rangle$ to $\langle \Phi \rangle$.

Let government i's payoff as a function of trade regime r be denoted by $g_i(r)$, where $r \in \{\langle \Phi \rangle, \langle ij \rangle, \langle ik \rangle, \langle jk \rangle, \langle ih \rangle, \langle jh \rangle, \langle kh \rangle, \langle F \rangle\}$. Also, let $\Delta g_i(r - v)$ denote the difference between government i's payoff under trade regimes r and v:

$$\Delta g_i(r - v) \equiv g_i(r) - g_i(v) \tag{10}$$

We next consider the tariff determination under the bilateralism game.

---

[17]Note that the FTA $\langle ij \rangle$ obtains so long as country i and j call only each other, regardless of the nature of country k's announcement. Thus, if $\sigma_i = \{j, \phi\}$ and $\sigma_j = \{i, \phi\}$, then country k would be indifferent between $\sigma_k = \{\phi, \phi\}$, $\{i, \phi\}$, $\{\phi, j\}$ and $\{i, j\}$ because its announcement has no bearing upon the outcome. In this case, we assume that country k makes the most parsimonious announcement among the three, $\sigma_k = \{\phi, \phi\}$. The intuitive justification for this assumption is that an FTA proposal is likely to be costly in the real world and a country that receives no proposals from others would be better off not making any proposals of its own.

3.1 Optimal tariffs

Since Article I of GATT forbids tariff discrimination, we assume that under No Agreement $\langle\Phi\rangle$, each country imposes a non-discriminatory tariff on its trading partners: $t_{ij} = t_{ik} = t_i(\Phi)$.[18] If country i is not part of any FTA, its optimal MFN tariff is calculated as

$$t_i(\Phi) \equiv \arg\max g_i(\Phi) = \frac{(2\beta_i + 3)(\alpha - c)}{2(5 - 2\beta_i)} \tag{11}$$

As expected, governments that are influenced heavily by special interest groups rely on more protectionist trade policies: $\frac{\partial t_i(\Phi)}{\partial \beta_i} > 0$. Note that for large values of β_i a prohibitive tariff obtains:

$$x_{ji}(\Phi) = \frac{(2\beta_i - 1)(\alpha - c)}{2(5 - 2\beta_i)} < 0 \text{ when } \beta_i > \frac{1}{2}$$

In order to exclude prohibitive tariffs and guarantee market access for all exporting firms, we assume $\beta_i \leq \frac{1}{2}$ holds for all i.[19]

If two countries form an FTA, they remove their tariffs on each other and impose optimal external tariffs on a non-member country: under $\langle ij\rangle$ we have $t_{ij} = t_{ji} = 0$, $t_{ki} = t_i(ij)$ and $t_{kj} = t_j(ij)$. The optimal external tariff of country i on the non-member country k is given by:

$$t_i(ij) \equiv \arg\max g_i(ij) = \frac{(2\beta_i + 3)(\alpha - c)}{21 - 2\beta_i} \tag{12}$$

Similar to $t_i(\Phi)$, the optimal external tariff under an FTA is increasing in the degree of political bias: $\frac{\partial t_i(ij)}{\partial \beta_i} > 0$. However, via reducing the markup for the domestic firm due to higher competition, an FTA reduces the effectiveness of the external tariff in extracting rents from the outside firms. Thus, the formation of a bilateral FTA induces each member to lower its tariff on the non-member country relative to the No Agreement $\langle\Phi\rangle$ and the model exhibits the tariff complementarity:[20]

$$t_i(\Phi) - t_i(ij) = \frac{(2\beta_i + 11)(2\beta_i + 3)(\alpha - c)}{2(21 - 2\beta_i)(5 - 2\beta_i)} > 0$$

Furthermore, the tariff complementarity effect deepens as the country becomes more politically motivated:[21]

---

[18] Note that, under no agreement, cost symmetry across firms implies that non-discrimination arises as an equilibrium policy in the model.

[19] This guarantees positive output levels of non-numeraire good under all possible agreements.

[20] See Bagwell and Staiger (1997) for a detailed discussion of the tariff complementarity effect and Estevadeordal. et. al (2007) for empirical evidence in its support. It is worth noting that tariff complementarity also arises in simple general equilibrium models of free trade agreements such as Bond et. al (2004).

[21] See Ornelas (2005b) for a detailed discussion.

$$\frac{\partial[t_i(\Phi) - t_i(ij)]}{\partial \beta_i} > 0$$

3.2 Equilibrium analysis under symmetry

In this section we derive equilibrium trade policy under symmetry in political prefer-ences. Therefore, throughout this section, as well as in Section 4, we maintain that countries share the same political bias parameter: $\beta_i = \beta$ *for all* i.[22]

3.2.1 Nash equilibria

Four trade policy regimes can potentially appear in the equilibrium – No Agreement $\langle \Phi \rangle$, a single bilateral FTA $\langle ij \rangle$, a hub and spoke agreement $\langle ih \rangle$, and a global free trade $\langle F \rangle$. It is straightforward to show that the announcements leading to No Agreement is a Nash equilibrium since no country has an incentive to announce another's name if the latter does not announce its name in return. Before we discuss which other announcement profiles can emerge as Nash equilibria, let $\beta(r - v)_i$ denote the critical threshold of the degree of political bias parameter at which government i is indifferent between regimes r and v.

    Lemma 1: *Under symmetry, the following holds for all* $i, j, k = a, b, c$:

    (i) $g_i(ih) > \max\{g_i(F), g_i(jh), g_i(ij), g_i(jk), g_i(\Phi)\}$ *while* $\Delta g_k(ih - ij) < 0$;

    (ii) $\Delta g_i(ij - \Phi) = \Delta g_j(ij - \Phi) \geq 0$ *iff* $\beta \leq \beta(ij - \Phi)_i$;

    (iii) $\Delta g_j(F - ih) = \Delta g_k(F - ih) \geq 0$ *iff* $\beta \leq \beta(F - ih)_j$ *and*

    (iv) $\Delta g_i(F - jk) \geq 0$ *iff* $\beta \leq \beta(F - jk)_i$.

    Lemma 1(i) implies that, in comparison to all other trade regimes, country i obtains the highest welfare as being the hub under the hub and spoke agreement $\langle ih \rangle$. It also states that each spoke country (say country k) has an incentive to unilaterally change its announcement that leads to a break up its FTA with the hub and deviate from $\langle ih \rangle$ to $\langle ij \rangle$. This immediately implies that the announcement profile leading to a hub and spoke arrangement $\langle ih \rangle$ is not a Nash equilibrium.

    Turning to the announcement profile leading to a bilateral FTA $\langle ij \rangle$, it is immediate from Lemma 1(ii) that a member country of a bilateral FTA has a unilateral incentive to change its announcement and break the agreement when its political bias is sufficiently strong: $\beta > \beta(ij - \Phi)_i$. This implies that the announcement profile leading to a bilateral FTA is a Nash equilibrium when $\beta \leq \beta(ij - \Phi)_i$ holds. It is also worth noting here that the tariff complementarity effect is large enough to make a non-member country always better off under a bilateral FTA relative to the status quo:

$$\Delta g_k(ij - \Phi) > 0 \text{ for all } \beta \tag{13}$$

The only remaining candidate for a Nash equilibrium is free trade $\langle F \rangle$. For the announcement profile leading to $\langle F \rangle$ to be a Nash equilibrium, we need to rule out the following two announcement deviations of country k: (1) unilateral deviation from announcement of $\{i, j\}$ to $\{i, \phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle ih \rangle$ and (2) unilateral deviation from announcement of $\{i, j\}$ to $\{\phi, \phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle ij \rangle$. We know from part (i) of Lemma 1 that a hub country is always better off under the hub and spoke agreement relative to free trade. Part (iii) of Lemma 1 implies that a country (say k) has an incentive to unilaterally deviate from $\{i, j\}$ to $\{i, \phi\}$ when the degree of political motivation is sufficiently high. We also know from part (iv) of Lemma 1 that a country prefers free riding on bilateral trade liberalization undertaken by others without liberalizing itself when it has strong political bias.

A comparison of the two critical values for the political bias parameter in parts (iii) and (iv) of Lemma 1 reveals that $\beta(F - jk)_i < \beta(F - ih)_j$. Thus, the announcement profile leading to $\langle F \rangle$ is a Nash equilibrium only when $\beta \leq \beta(F - jk)_i$ holds. The above discussion yields that we have multiple Nash equilibria depending on the values of β. To deal with the multiplicity of equilibria described above and to capture the process of FTA formation in a more realistic fashion, we now isolate coalition proof Nash equilibria (CPNE), i.e. the Nash equilibria that are immune to self-enforcing coalitional deviations.[23]

In order to verify if free trade $\langle F \rangle$ is a CPNE, where $\sigma_i = \{j, k\}$, $\sigma_j = \{i, k\}$, and $\sigma_k = \{i, j\}$, we need to rule out two distinct joint deviations:

(J1) joint deviation of countries i and j to a pair of announcements where they only name each other with the intention of exuding the third country and forming an FTA: country i changes its announcement from $\{j, k\}$ to $\{j, \phi\}$ and country j changes its announcement from $\{i, k\}$ to $\{i, \phi\}$ while country k's announcement stays fixed at $\{i, j\}$;

(J2) joint deviation of two or all three countries to announcements that lead to a deviation from $\langle F \rangle$ to $\langle \Phi \rangle$.

First note that when countries i and j have sufficiently high political motivation, they have an incentive to exclude country k and (J1) occurs:

$$\Delta g_i(F - ij) = \Delta g_j(F - ij) \geq 0 \text{ iff } \beta \leq \beta(F - ij)_i \tag{14}$$

However, note from Lemma 1 that such joint deviation is not self-enforcing because, holding country k's announcement *fixed* at $\sigma_k = \{i, j\}$, each member of the deviating

---

[23]An alternative approach would be to use the notion of a strong Nash equilibrium (SNE). However, the use of CPNE is more appealing since the definition of a SNE requires that the equilibrium strategies be immune to *any* joint deviations, even those that are not self-enforcing (i.e. are susceptible to further deviations on the part of a proper subset of the deviating coalition).

coalition (comprised of countries i and j) has an incentive to *further deviate* by altering its announcement to also name country k. To see why this is the case, suppose country i further alters its announcement from $\{j, \phi\}$ to $\{j, k\}$ given that country j announces $\{i, \phi\}$ and country k's announcement is *fixed* at $\{i, j\}$. This further deviation on the part of country i is payoff-improving because it leads to the hub and spoke arrangement $\langle ih \rangle$ under which it is better off relative to free trade (because of its status as the hub country). Since this further deviation of country i from announcing $\{j, \phi\}$ to $\{j, k\}$ is credible, the original joint deviation (J1) is *not self-enforcing.*

Now consider the joint deviation (J2). Countries prefer free trade to no agreement when their valuation of national welfare relative to political contributions is high enough:

$$\Delta g_i(F - \Phi) \geq 0 \ iff \ \beta \leq \beta(F - \Phi)_i \tag{15}$$

Thus (J2) occurs when $\beta > \beta(F - \Phi)_i$ holds. Note also that $\beta(ij - \Phi)_i > \beta(F - \Phi)_i$ holds and thus (J2) is self-enforcing only when $\beta \geq \beta(ij - \Phi)_i$. Given that $\beta(ij - \Phi)_i > \beta(F - jk)_i$, the range over which the strategy profile leading to free trade $\langle F \rangle$ is CPNE is determined by the unilateral deviation of country i from announcement of $\{j, k\}$ to $\{\phi, \phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle jk \rangle$. As a result, free trade is a CPNE when $\beta \leq \beta(F - jk)_i$ holds.[24]

Next, we consider whether the Nash equilibrium announcement profile leading to a bilateral FTA $\langle ij \rangle$ is a CPNE. To this end, we examine three possible coalitional announcement deviations: deviation of countries i, j and k from their respective announcements $\{j, \phi\}$, $\{i, \phi\}$ and $\{\phi, \phi\}$ to the announcements $\{j, k\}$, $\{i, k\}$ and $\{i, j\}$, leading to a deviation from $\langle ij \rangle$ to $\langle F \rangle$ (JB1); deviation of countries i and k from their respective announcements $\{j, \phi\}$ and $\{\phi, \phi\}$ to $\{j, k\}$ and $\{i, \phi\}$, holding country j's announcement *fixed* at $\sigma_j = \{\{i, \phi\}\}$, leading to a deviation from $\langle ij \rangle$ to $\langle ih \rangle$ (JB2); and deviation of countries i, j and k from their respective announcements $\{j, \phi\}$, $\{i, \phi\}$, and $\{\phi, \phi\}$ to $\{\phi, k\}$, $\{\phi, k\}$, $\{i, j\}$, leading to a deviation from $\langle ij \rangle$ to $\langle kh \rangle$ (JB3).

It is immediate from Lemma 1(iv) and (14) that JB1 happens when $\beta < \beta(F - jk)_i$ and it is a self-enforcing deviation since $\langle F \rangle$ is a strong Nash equilibrium over this range. We know from Lemma 1(i) that JB2 never occurs since country k has no incentive to deviate from its announcement $\{\phi, \phi\}$ to $\{i, \phi\}$, leading to a deviation $\langle ij \rangle$ to $\langle ih \rangle$ (joint with the deviation of country i from its announcement $\{j, \phi\}$ to $\{j, k\}$). Similarly, Lemma 1(i) implies that even when JB3 happens, it is not self enforcing since either country i or country j has a unilateral incentive to break up their FTA with the hub country. This discussion with the above Nash condition informs us that a bilateral FTA is a CPNE when $\beta(ij - \Phi)_i \geq \beta \geq \beta(F - jk)_i$.

---

[24]In fact, free trade is also a strong Nash equilibrium when $\beta \leq \beta(F - jk)_i$ since there exist no unilateral or coalitional deviation from free trade.

Finally consider no agreement $\langle \Phi \rangle$. We know from Lemma1(ii) that the joint deviation of countries i and j from their respective announcements $\{\phi, \phi\}$ and $\{\phi, \phi\}$ to $\{j, \phi\}$ and $\{i, \phi\}$, leading to a deviation from $\langle \Phi \rangle$ to $\langle ij \rangle$, happens when $\beta < \beta(ij - \Phi)_i$ and it is self enforcing. We show in Appendix A that when $\beta \geq \beta(ij - \Phi)_i$ holds, there exists no other coalitional deviation from the Nash equilibrium strategy profile leading to No Agreement $\langle \Phi \rangle$, and thus $\langle \Phi \rangle$ is stable.

Proposition 1: *Given symmetry, $\beta(ij - \Phi)_i > \beta(F - jk)_i$ holds and the following results obtain under bilateralism:*

(i) free trade $\langle F \rangle$ is the unique CPNE when $\beta < \beta(F - jk)_i$;

(ii) bilateral FTAs $\langle ij \rangle$ are CPNE when $\beta(ij - \Phi)_i \geq \beta \geq \beta(F - jk)_i$ and

(iii) No Agreement $\langle \Phi \rangle$ is the unique CPNE when $\beta > \beta(ij - \Phi)_i$.

The above proposition informs us that even when countries have completely symmetric political preferences, multilateral free trade fails to be a CPNE when government's valuation of producer surplus is high relative to other welfare components. This result points to the importance of political economy factors even under symmetry and it deviates from Saggi and Yildiz (2010) who argue in an endowment model that bilateralism leads to global free trade when countries are completely symmetric with respect to their endowment levels. It is worth noting that when the degree of political motivation is sufficiently low, the above result implies that free trade is a unique stable agreement under bilateralism. This result is similar to Goyal and Joshi (2006) who also employ an oligopoly model of trade but assume $\beta = 0$. As we shall see below, by comparing equilibria obtained under bilateralism and multilateralism, we provide an alternative interpretation of this result relative to Goyal and Joshi (2006).

4   Endogenous agreements under multilateralism

In this section we analyze equilibrium trade policy formation under a scenario when countries cannot pursue preferential trade liberalization and can only reduce tariffs multilaterally on an MFN basis. We refer to this scenario as *multilateralism*. Comparison of the equilibria that arise under bilateralism and multilateralism would reveal whether the freedom to pursue FTAs, incorporated in the GATT Article XXIV, results in deeper trade liberalization or not.

Under a multilateral approach to trade liberalization, the strategy set of country i is $\Omega_i = \{\phi, M\}$. In other words, each country can announce either in favor of or against multilateral tariff reductions. If all three countries announce in favor, they choose to set optimal tariffs in order to maximize joint welfare. Such tariff is equal to zero in our model. If only a pair of countries i and j announce in favor of multilateralism, they choose optimal tariffs in order to maximize their joint welfare subject to the constraint

that they cannot discriminate against country k – i.e. in accordance with the MFN clause of the WTO, the tariffs that they impose on each other must be equal to their respective tariffs on country k. Formally, countries i and j sign the multilateral agreement $\langle ij^m \rangle$ when individual country announcements are as follows: $\sigma_i = M$, $\sigma_j = M$, $\sigma_k = \phi$. Finally, we should note that if two (or more) countries announce against multilateralism, No Agreement $\langle \Phi \rangle$ prevails under which each country imposes its optimal MFN tariff on every other country.

Saggi and Yildiz (2011) use an alternative model for the multilateralism game where two countries are not permitted to undertake trade liberalization on an MFN basis and a multilateral agreement could only arise if all three countries participated in it. This 'unanimity' view of multilateralism surely captures some aspects of the WTO but we think that our model is more general and captures multilateral trade liberalization more accurately. More specifically, following Saggi and Yildiz (2010), our model assumes that if two countries announce in favor of multilateralism and the third one announces against it, then the two that are in favor can lower their tariffs on each other as long as they extend the tariff reductions to the third country. In other words, the model allows two countries to liberalize trade on an MFN basis without needing the third country's consent. It is important to note that modeling multilateralism in this way also allows us to compare the degree of trade liberalization under multilateralism and bilateralism, something that could not be addressed under the "all or nothing" formulation used in Saggi and Yildiz (2011).

4.1 Optimal tariffs

As indicated above, when countries i and j agree to sign a multilateral agreement $\langle ij^m \rangle$ they jointly choose tariffs (t_i^m, t_j^m) to solve

$$[t_i(ij^m), t_j(ij^m)] \equiv \arg\max \left[g_i(ij^m) + g_j(ij^m) \right] \tag{16}$$

which leads to the following optimal MFN tariffs:

$$t_z^m = \frac{(2\beta_z + 1)(\alpha - c)}{2(3 - 2\beta_z)}, \ z = i, j \tag{17}$$

First we find that, similar to $t_z(ij)$ under $\langle ij \rangle$, $t_z(ij^m)$ also increases in the degree of political motivation: $\frac{\partial t_z(ij^m)}{\partial \beta_z} > 0$. Note also that countries that sign the multilateral agreement $\langle ij^m \rangle$ lower their tariffs on each other as well as on the non-participating country (i.e. k) relative to the No Agreement $\langle \Phi \rangle$:

$$t_z(\Phi) - t_z(ij^m) = \frac{2(1 - 2\beta_z)(\alpha - c)}{(3 - 2\beta_z)(5 - 2\beta_z)} > 0, \ z = i, j \tag{18}$$

The inequality $t_z(ij^m) < t_z(\Phi)$ captures the *free-riding* problem inherent to multilateral trade liberalization when it does not involve all three countries – under $\langle ij^m \rangle$ country k benefits from the multilateral trade liberalization undertaken by the other two countries without having to offer any liberalization in return since it retains its optimal Nash tariff $t_k(\Phi)$ on countries i and j.

Furthermore, country k faces lower tariffs in export markets when the other two countries implement a preferential rather than a multilateral agreement, i.e., $t_z(ij) < t_z(ij^m)$, $z = i, j$. In other words, the degree of trade liberalization undertaken by two countries is lower when they sign a multilateral agreement $\langle ij^m \rangle$ relative to the one when they sign an FTA $\langle ij \rangle$: $t_z(\Phi) - t_z(ij^m) < t_z(\Phi) - t_z(ij)$, $z = i, j$. Note also that the difference between $t_z(ij)$ and $t_z(ij^m)$ widens as the countries becomes more politically motivated: $\frac{\partial [t_z(ij^m) - t_z(ij)]}{\partial \beta_z} > 0$.

From a non-member country's perspective (i.e. country k), the relative effect of a preferential and a multilateral agreements between i and j on country k's welfare is explained by two distinct effects. First, there is a *tariff level effect*, whereby country k faces different tariff levels under FTA and multilateral agreement between i and j. Since country k faces lower tariff under $\langle ij \rangle$ relative to $\langle ij^m \rangle$, tariff level effect is stronger under bilateralism. Second, while the non-member country faces symmetric treatment under $\langle ij^m \rangle$, there is a *discrimination effect* stemming from the fact that the non-member country is subject to a discriminatory treatment in each member country's market under the bilateral FTA $\langle ij \rangle$: while countries i and j face zero tariffs in each other's market under $\langle ij \rangle$, country k faces the tariff $t_z(ij)$, $z = i, j$. Since there is no discrimination under $\langle ij^m \rangle$, discrimination effect implies additional welfare loss for country k under bilateralism. It turns out that the welfare (and the government's payoff) of the non-member country (k) is lower under $\langle ij \rangle$ compared to $\langle ij^m \rangle$ when β is sufficiently low:

$$g_k(ij) \leq g_k(ij^m) \text{ when } \beta_z \leq \beta(ij - ij^m)_k, \ z = i, j \tag{19}$$

When governments i and j have low political biases (β_i and β_j are sufficiently low), $t_z(ij)$ and $t_z(ij^m)$ are relatively close and tariff complementarity effect is relatively small. Thus, under such a case, the discrimination effect under bilateralism dominates the tariff level effect under multilateralism and the non-member country prefers $\langle ij^m \rangle$ to $\langle ij \rangle$. As β_i and β_j rise, the tariff complementarity deepens and the difference between $t_z(ij)$ and $t_z(ij^m)$ rises. Therefore, the tariff level effect becomes more pronounced relative to the discrimination effect so that the optimal strategy of the non-member country gets reversed. As we demonstrate below, this fundamental difference between preferential and multilateral trade liberalization plays a crucial role in our analysis.

4.2 Equilibrium analysis under symmetry

In this section, we derive the set of CPNE of the multilateralism game for the case of complete symmetry in the political economy preferences between countries: $\beta_i = \beta$ for all i. As under bilateralism, the status quo $\langle \Phi \rangle$ is a Nash equilibrium under multilateralism. In order to check whether the announcement profile leading to $\langle ij^m \rangle$ is also a Nash equilibrium, we need to consider two unilateral deviations: (U1) - unilateral deviation of country i from its announcement of $\{M\}$ to the announcement $\{\phi\}$ that leads to a deviation from $\langle ij^m \rangle$ to $\langle \Phi \rangle$; (U2) - unilateral deviation of country k from its announcement of $\{\phi\}$ to the announcement $\{M\}$ that leads to a deviation from $\langle ij^m \rangle$ to $\langle F \rangle$

It is easy to show that a member country has no incentive to break the multilateral agreement $\langle ij^m \rangle$:

$$\Delta g_i(ij^m - \Phi) = \Delta g_j(ij^m - \Phi) \geq 0 \text{ for all } \beta \tag{20}$$

while the outsider country (k) benefits from joining the agreement $\langle ij^m \rangle$, thereby converting it to $\langle F \rangle$, only when the degree of political bias is sufficiently low:

$$\Delta g_k(F - ij^m) > 0 \text{ iff } \beta < \beta(F - ij^m)_k \tag{21}$$

Thus, under symmetry the announcement profile leading to the multilateral agreement $\langle ij^m \rangle$ is a Nash equilibrium if $\beta \geq \beta(F - ij^m)_k$ holds. This result is interesting because it says that when β is sufficiently large, the free rider benefits from the multilateral trade liberalization undertaken by other countries are large enough for country k to make it opt out of $\langle F \rangle$ and thereby prevent global free trade from emerging. The only remaining question is whether the announcement profile leading to free trade $\langle F \rangle$ is a Nash equilibrium. The answer is in the affirmative: the only possible unilateral deviation that can occur from free trade is the unilateral deviation of country k from its announcement of $\{M\}$ to the announcement $\{\phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle ij^m \rangle$. We know from (21) that this deviation occurs when $\beta > \beta(F - ij^m)_k$. Thus, the announcement profile leading to free trade $\langle F \rangle$ is a Nash equilibrium when $\beta \leq \beta(F - ij^m)_k$.

Note that, as before, multiple Nash equilibria arise in the multilateralism game – the key difference being that, by definition, a bilateral FTA is infeasible under the multilateralism game. We next isolate CPNE under the two approaches to trade liberalization. It is clear from (20) that the No Agreement $\langle \Phi \rangle$ fails to be a CPNE since any two countries have incentives to jointly deviate from their announcements of no agreement to announcements of multilateralism that leads to a deviation from $\langle \Phi \rangle$ to $\langle ij^m \rangle$ and it is a self-enforcing deviation. We thus have

Proposition 2: *Given symmetry, the following holds under multilateralism:*

(i) free trade $\langle F \rangle$ is the unique CPNE when $\beta < \beta(F - ij^m)_k$ and

(ii) a multilateral agreement $\langle ij^m \rangle$ is the unique CPNE when $\beta \geq \beta(F - ij^m)_k$.[25]

4.2.1 Comparison of stable agreements under bilateralism and multilateralism

A comparison of Propositions 1 and 2 allows us to assess the role of GATT Article XXIV on the world trading system and leads to our main result under symmetry, depicted in Figure 1:

Proposition 3: *Given symmetry, $\beta(ij - \Phi)_i > \beta(F - ij^m)_k > \beta(F - ij)_k$ holds and the following results obtain:*

(i) (Irrelevance of Bilateralism) when $\beta < \beta(F - ij)_k$ holds, $\langle F \rangle$ is the the unique CPNE under both bilateralism and multilateralism.

(ii) (Strong Stumbling Block) when $\beta(F - ij)_k < \beta < \beta(F - ij^m)_k$ holds, the unique CPNE under multilateralism is $\langle F \rangle$ whereas under bilateralism it is $\langle ij \rangle$;

(iii) (Partial Building Block) when $\beta(F - ij^m)_k < \beta < \beta(ij - \Phi)_i$ the unique CPNE under multilateralism is $\langle ij^m \rangle$ whereas under bilateralism it is $\langle ij \rangle$ and

(iv) (Partial Stumbling Block) when $\beta > \beta(ij - \Phi)_i$ holds, the unique CPNE under multilateralism is $\langle ij^m \rangle$ whereas under bilateralism it is $\langle \Phi \rangle$.

– Figure 1 –

The first part of the above proposition implies that when countries are completely symmetric, the freedom to pursue bilateral FTAs does not matter only when β is sufficiently low. Therefore, when governments have low political bias, nothing would be lost by forsaking the freedom to pursue FTAs since such agreements would not even arise in equilibrium. However, it follows from the second part of the proposition that when β is at the intermediate range, the ability to form FTAs reduces the likelihood of global free trade arising as a CPNE. When such is the case, we argue that FTAs act as *strong stumbling blocs* for multilateral free trade. The intuition for this result can be explained as follows. First note that when $\beta(F - ij)_k < \beta < \beta(F - ij^m)_k$ holds, it is immediate from (19) that $\beta > \beta(ij - ij^m)_k$ obtains and the *level effect* under multilateralism dominates the *discrimination effect* under bilateralism.[26] Therefore, country k has an incentive to free ride on the trade liberalization by countries i and j under bilateralism by deviating unilaterally from $\langle F \rangle$ to $\langle ij \rangle$, while such incentive does not exist under multilateralism.

When the global free trade fails to obtain under both bilateralism and multilateralism, the third part of the proposition argues that for $\beta(F - ij^m)_k < \beta < \beta(ij - \Phi)_i$,

---

[25] In fact, we can show that under symmetry $\langle F \rangle$ and $\langle ij^m \rangle$ are the strong Nash equilibrium over the given parametre ranges.

[26] Note that under symmetry $\beta_z(ij - ij^m)_k = \beta(ij - ij^m)_k$, $z = i, j$.

the ability to form bilateral FTAs can improve global welfare since the degree of trade liberalization undertaken by two countries is lower when they sign a multilateral agreement $\langle ij^m \rangle$ relative to the case when they sign a bilateral FTA $\langle ij \rangle$. In such a case, we argue that FTAs act as *partial building blocs*. Finally, when β is sufficiently large (i.e. tariffs are close to prohibitive), the ability to form FTAs reduces the world welfare (FTAs act as *partial stumbling blocs*) because partial multilateral trade liberalization is always better than no trade liberalization under bilateralism.

Given these results under symmetry, it is natural to ask: does the freedom to pursue bilateral FTAs ever help to achieve global free trade? We next show that such possibility arises when we introduce asymmetry in political preferences between countries.

5   Equilibrium analysis under asymmetry

Hereafter, we drop the assumption that countries have symmetric political preferences. Before proceeding with the equilibrium analysis of trade policies, we first investigate how individual country's incentives to form a bilateral FTA depend on the degree of asymmetry in political preferences across countries. We address this key question by breaking it up into parts and stating two related lemmas.

Lemma 2: *Let the initial regime be denoted by v and countries i and j form an FTA that leads to regime r. Then, for any initial regime v, the following holds:*

(i) $\frac{\partial \Delta g_i(r-v)}{\partial \beta_i} < 0$; *(ii)* $\frac{\partial \Delta g_i(r-v)}{\partial \beta_j} > 0$ and *(iii)* $\frac{\partial \Delta g_i(r-v)}{\partial \beta_k} = 0$

Joining an FTA involves a trade-off for member countries. On one hand, an FTA is costly for member countries because their domestic surpluses are lowered under regime r relative to the regime v. Part (i) of Lemma 2 states that the reduction in domestic surplus, implied by an FTA, is increasing in the degree of a country's own political bias. On the other hand, being part of an FTA increases export profits of members in each other's markets, which is increasing in the external tariff and thus in the degree of political bias in the partner country (part (ii) of Lemma 2). The third part simply follows from the market segmentation: export profit of countries i and j in country k are the same under both regimes r and v.

Next, we consider the effect of an FTA between countries i and j on the government k's payoff.

Lemma 3: *Let the initial regime be denoted by v. If countries i and j form an FTA that leads to regime r, then the following holds for any initial regime v:*

(i) $\frac{\partial \Delta g_k(r-v)}{\partial \beta_z} < 0$ *when z is an FTA partner of country k in both regimes r and v, $z = i, j$.*

(ii) $\frac{\partial \Delta g_k(r-v)}{\partial \beta_z} \geq 0$ *if and only if $\beta_z \leq \overline{\beta}$ when z is not an FTA partner of country k in regimes r and v, $z = i, j$.*

(iii) $\frac{\partial \Delta g_k(r-v)}{\partial \beta_k} = 0$

To see the intuition behind the first part of the above lemma, consider the case when the status-quo regime v is $\langle ik \rangle$ and i and j forms an FTA so that the new regime r is $\langle ih \rangle$. Under $\langle ik \rangle$, country k's firm enjoys a strategic advantage against country j's firm in country i's market. This strategic advantage gets larger as β_i increases since country j's firm faces higher tariffs in country i's market. With an additional FTA between countries i and j under $\langle ih \rangle$ (regime r), this strategic advantage in country i's market disappears. Thus, the first part of the lemma obtains.[27] To see the intuition for the second part, consider the case when the status-quo regime v is $\langle jk \rangle$ and i and j forms an FTA so that the new regime r is $\langle jh \rangle$. Note that country k faces a lower tariff in country i's market under $\langle jh \rangle$ relative to the one under $\langle jk \rangle$ due to the tariff complementarity effect. We know that the tariff complementarity effect (*tariff level effect*) rises in β_i: $\frac{\partial [t_i(\Phi) - t_i(ij)]}{\partial \beta_i} > 0$. However, under $\langle jh \rangle$, lower tariff comes in a discriminatory way since country j faces zero tariff in country i's market while country k faces a positive one (*negative discrimination effect*), and the discrimination effect rises in β_i: $\frac{\partial t_i(ij)}{\partial \beta_i} > 0$. As a result, relative importance of the two effects determines how $g_k(jh - jk)$ changes with β_i. While the tariff level effect dominates the discrimination effect when β_i is sufficiently low $(\beta_i \leq \overline{\beta})$, it gets reversed when the degree of political motivation is sufficiently high $(\beta_i > \overline{\beta})$.[28] Finally, as in Lemma 2, part *(iii)* of Lemma 3 follows directly from the market segmentation assumption.

It is important to note that $g_i(ih) > g_i(F)$ holds not only under symmetry but also when countries have asymmetric political biases. First note that the hub country i enjoys privileged access in both foreign countries under $\langle ih \rangle$ – neither spoke country imposes a tariff on the hub country whereas both impose external tariffs on each other. As a result of this favorable treatment, country i's export surplus under $\langle ih \rangle$ exceeds that under $\langle F \rangle$. Second, country i's domestic surplus under $\langle ih \rangle$ is the same as that under $\langle F \rangle$. Thus, country i is strictly better off under $\langle ih \rangle$ relative to $\langle F \rangle$. Note from part (i) of Lemma 3 that the positive effect of the privileged access gets larger as the spoke countries are more politically motivated and have more protected domestic markets.

Finally, we examine how the incentives of countries to form (or to join) a multilateral agreement depend on the degree of asymmetry in the political preference across countries.

Lemma 4: *Under multilateralism, the following holds:*
(i) $\frac{\partial \Delta g_i(ij^m - \phi)}{\partial \beta_i} \geq 0$ *while* $\frac{\partial \Delta g_i(ij^m - \phi)}{\partial \beta_j} \leq 0$ *and* $\frac{\partial \Delta g_i(ij^m - \phi)}{\partial \beta_k} = 0$;
(ii) $\frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_i} < 0$ *while* $\frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_j} \geq 0$ *and* $\frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_k} \geq 0$

---

[27] The same intuition and result obtain when the status-quo regime v is $\langle kh \rangle$ and i and j forms an FTA so that the new regime r is $\langle F \rangle$.

[28] The same intuition and result obtain when the status-quo regime v is $\langle \Phi \rangle$ and i and j forms an FTA so that the new regime r is $\langle ij \rangle$.

The intuition behind the results reported in Lemma 4 is analogous to that obtained for bilateralism in Lemmas 2 and 3 with one exception: whereas $\frac{\partial \Delta g_i(ij-\phi)}{\partial \beta_i} < 0$ and $\frac{\partial \Delta g_i(ij-\phi)}{\partial \beta_j} > 0$ hold under bilateralism, the opposite is true under multilateralism, i.e., $\frac{\partial \Delta g_i(ij^m-\phi)}{\partial \beta_i} > 0$ and $\frac{\partial \Delta g_i(ij^m-\phi)}{\partial \beta_j} < 0$. To see why this is the case, recall that under the multilateral agreement $\langle ij \rangle^m$, countries i and j impose jointly chosen MFN tariffs not only on country k but also on each other, while under the bilateral agreement $\langle ij \rangle$ they lower the internal tariffs to zero. Since country i's ability to protect its own market rises in β_i while the export surplus gain in country j's market falls in β_j, the above difference arises.

5.1 Equilibrium agreements under bilateralism

Let the pattern of asymmetry in political motivations to be given by: $\beta_b = \beta_c = \beta \leq \beta_a \leq \frac{1}{2}$.[29] Thus, country a is more concerned about producer surplus than countries c and b. Also, let $\beta_a(r-v)_i$ denote the critical threshold for country a's political bias parameter at which government i is indifferent between regimes r and v. To avoid redundancy, we skip the discussion of Nash equilibrium and focus directly on CPNE agreements under bilateralism. Before proceeding, it is useful to establish the following result that generalizes Lemma 1 from the perspective of countries with lower political biases:

Lemma 5: *When $\beta_b = \beta_c = \beta \leq \beta_a \leq \frac{1}{2}$ holds, $g_i(ih) > \max\{g_i(F), g_i(jh), g_i(ij), g_i(jk), g_i(\Phi)\}$ for all $i, j = b, c$.*

The above lemma is a direct implication of the combination of Lemmas 1, 2, and 3, and implies that, in comparison to all other trade agreements, countries with lower political biases obtain the highest welfare as being the hub under the hub and spoke agreement $\langle ih \rangle$. Note that Lemma 5 does not apply to the country with the largest political bias since it has stronger incentive to break its FTA link(s).[30]

The following proposition summarizes which trade regimes can arise in the equilibrium under bilateralism and asymmetry in political preferences between countries.

Proposition 4: *Suppose $\beta_b = \beta_c = \beta \leq \beta_a \leq \frac{1}{2}$. Then, the following holds under bilateralism:*

(i) free trade $\langle F \rangle$ is the unique CPNE when $\beta_a < \beta_a(F - bc)_a$;

---

[29] In Appendix B, we show that all results obtained in this section remain qualitatively the same under an alternative pattern of asymmetry when two countries have stronger political biases than the third one: $\beta_a \leq \beta_b = \beta_c = \beta \leq \frac{1}{2}$.

[30] Since countries b and c are identical, the same payoff obtains for country a under $\langle ab \rangle$ and $\langle ac \rangle$ (and under $\langle bh \rangle$ and $\langle ch \rangle$). Similarly, country b's payoff under $\langle ab \rangle$ and $\langle bh \rangle$ is the same as country c's payoff under $\langle ac \rangle$ and $\langle ch \rangle$, respectively. From heron, for notational simplicity, we opt for using $\langle ab \rangle$ to represent a bilateral FTA between the country with the largest political bias and the one with lower political bias and $\langle bh \rangle$ to represent hub and spoke regime where the hub country has a lower political bias. In a similar way, under multilateralism game, $\langle ab^m \rangle$ represents a two-country multilateral agreement between the country with the largest political bias and the one with lower political bias.

(ii) any bilateral FTA is a CPNE when $\min\{\beta_a(ah - ab)_c, \beta_a(ab - \Phi)_a\} \geq \beta_a \geq \beta_a(F - bc)_a;$

(iii) bilateral FTA $\langle bc \rangle$ *is the unique CPNE when (a)* $\beta_a > \beta_a(F - bc)_a$ *and (b)* $\beta_a > \min\{\beta_a(ah - ab)_c, \beta_a(ab - \Phi)_a\}\}$ *and (c)* $\beta < \beta(bc - \Phi)_b$ *holds and*

(iv) no agreement $\langle \Phi \rangle$ *is the unique CPNE when* $\beta > \beta(bc - \Phi)_b.$

– Figure 2 –

Proposition 4 relates the degree of underlying asymmetry in political preferences to the nature of agreements that are CPNE and is illustrated in Figure 2. Part (i) states that if both the degree of political preferences of countries and the degree of asymmetry in political preferences are sufficiently small, free trade is a unique CPNE. Recall that, under symmetric political preferences, the range of β for which free trade $\langle F \rangle$ obtains as a CPNE is determined by the unilateral deviation of country i from announcement of $\{j, k\}$ to $\{\phi, \phi\}$, which leads to a deviation from $\langle F \rangle$ to $\langle jk \rangle$. Under asymmetry, it is immediate from Lemma 2 that the stability of global free trade depends critically upon the unilateral deviation of the country with the strongest political bias (i.e. country a) from announcement $\{b, c\}$ to $\{\phi, \phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle bc \rangle$, and thus free trade $\langle F \rangle$ is a CPNE when $\beta_a \leq \beta_a(F - bc)_a$. Therefore, for free trade to be a CPNE, complete symmetry is not a necessary condition. As long as the degree of asymmetry in political preferences is sufficiently small, free trade is still a CPNE.

Part (ii) says that if the degree of asymmetry in political preferences is moderate, all bilateral trade agreements, including the country with the strongest political preference as a member, are CPNE. As the degree of political asymmetry rises, country a opts for not being a member of any FTA and a bilateral FTA between the other two countries is a unique CPNE. Finally part (iv) says that if the political preferences are strong, no agreement $\langle \Phi \rangle$ is a unique CPNE. It is noteworthy that multiple CPNE obtain when the degree of political preferences and political asymmetry are moderate – i.e. when $\min\{\beta_a(ah - ab)_c, \beta_a(ab - \Phi)_a\} \geq \beta_a \geq \beta_a(F - bc)_a$. The theory offers no guidance about which of these equilibria might be observed.

To see how the Article XXIV of the GATT matters for the world trade system, we next consider a scenario where countries were to follow only a multilateral approach to trade liberalization.

5.2 Equilibrium agreements under multilateralism

Before we identify agreements that are CPNE under a purely multilateral approach, it is important to note that, unlike in the bilateralism game, No Agreement $\langle \Phi \rangle$ never

arises as a CPNE. It follows from the inequality (20) and Lemma 4 that countries b and c have incentives to jointly deviate from their announcements of no agreement to announcements of multilateralism, and this deviation from $\langle \Phi \rangle$ to $\langle bc^m \rangle$ is self enforcing. In Proposition 5 we show that CPNE trade agreements under multilateralism depend on the strength of political preferences and the degree of asymmetry.

Proposition 5: *Suppose $\beta_b = \beta_c = \beta \leq \beta_a \leq \frac{1}{2}$. Then, the following holds under multilateralism:*

(i) free trade $\langle F \rangle$ is the unique CPNE when $\beta_a < \beta_a(F - bc^m)_a$;

(ii) any two-country multilateral agreement is a CPNE when $\beta_a \leq \min\{\beta_a(ab^m - \Phi)_b, \beta_a(F - ab^m)_c\}$ and

(iii) multilateral agreement $\langle bc^m \rangle$ is the unique CPNE when (a) $\beta_a > \beta_a(F - bc^m)_a$ and (b) $\beta_a > \min\{\beta_a(ab^m - \Phi)_b, \beta_a(F - ab^m)_c\}$ hold.

– Figure 3 –

Figure 3 illustrates Proposition 5 and shows which trade agreements are CPNE under multilateralism. The first part of the proposition states that when both the degree of political preferences of countries and the degree of asymmetry in political preferences are sufficiently small, free trade is the unique CPNE. Similar to our analysis under the bilateralism game, the unilateral deviation of country a from its announcement of $\{M\}$ to the announcement $\{\phi\}$ (leading to a deviation from $\langle F \rangle$ to $\langle bc^m \rangle$) determines the range over which $\langle F \rangle$ is a CPNE. Thus, free trade $\langle F \rangle$ is a CPNE of the multilateralism game when $\beta_a \leq \beta_a(F - bc^m)_a$. If this condition fails, the benefits of a free rider a from the multilateral trade liberalization undertaken by other countries are large enough for country a to opt out of free trade $\langle F \rangle$. In fact, free trade $\langle F \rangle$ is the strong Nash equilibrium when $\beta_a < \beta_a(F - bc^m)_a$ holds.

Proposition 5(ii) informs us that if countries have sufficient political preference and the degree of asymmetry in political preferences is moderate, all MFN-based trade agreement between any pair of countries are CPNE. As the degree of political preferences and political asymmetry rise, country a opts for not being a part of any MFN-based agreement and thus $\langle bc^m \rangle$ arises as a unique CPNE.

5.3   Does the ability to form bilateral FTAs matter?

We know from Propositions 4 and 5 that the global free trade is a CPNE under the bilateralism game for $\beta_a \leq \beta(F - bc)_a$ and it is a CPNE under multilateralism for $\beta_a \leq \beta_a(F - bc^m)_a$. A comparison of these critical threshold for political preferences leads to the following result:

Proposition 6: *Suppose $\beta_b = \beta_c = \beta \le \beta_a \le \frac{1}{2}$. Then, the following obtains:*

(i) (Irrelevance of Bilateralism) when $\beta_a < \min\{\beta_a(F - bc^m)_a, \beta_a(F - bc)_a\}$ holds, $\langle F \rangle$ is a unique CPNE under both bilateralism and multilateralism.

(ii) (Strong Building Block) when $\beta_a(F - bc^m)_a < \beta_a < \beta_a(F - bc)_a$ holds, a unique CPNE is $\langle F \rangle$ under bilateralism and $\langle bc^m \rangle$ under multilateralism;

(iii) (Strong Stumbling Block) when $\beta_a(F - bc)_a < \beta_a < \beta_a(F - bc^m)_a$ holds, a unique CPNE under multilateralism is $\langle F \rangle$ whereas under bilateralism $\langle ij \rangle$ is a CPNE, $\forall i, j$;

– Figure 4 –

Proposition 6 is the main result of this paper and is illustrated in Figure 4. It emphasizes the important role of asymmetry in political preferences between countries for the effect of the GATT Article XXIV on the global trade system. The first part of the above proposition implies that when both the degree of political preferences of countries and the degree of asymmetry in political preferences are sufficiently small, bilateralism is irrelevant for the ultimate objective of achieving global free trade. It is important to emphasize that whether free trade is a CPNE or not depends critically on the unilateral deviation incentive of the country with the strongest political preference (country a). Therefore, the comparison of welfare of country a under $\langle bc^m \rangle$ and $\langle bc \rangle$ is the key for the next two parts of the proposition. As discussed above, this comparison depends on the relative strengths of the tariff level effect under multilateralism and the discrimination effect under bilateralism. Recall that while country a faces higher tariff under $\langle bc^m \rangle$ than under $\langle bc \rangle$, it is discriminated against its rival exporter in each member country's market under the bilateral FTA $\langle bc \rangle$ but not under $\langle bc^m \rangle$.

The second part of the proposition informs us that, unlike in the symmetric case, Article XXIV *can be necessary* for achieving global free trade when the asymmetry in political preferences is sufficiently large and the political preferences of countries b and c are not too strong. This result stems from the fact that since political preferences of countries b and c are not too strong, country a faces lower MFN tariffs under $\langle bc^m \rangle$ and, despite the tariff complementarity effect under $\langle bc \rangle$, the discrimination effect dominates the tariff level effect. Thus, country a has less incentive to unilaterally deviate from its free trade announcement under bilateralism relative to multilateralism. As a result, when the asymmetry in political preferences is sufficiently large and the political preferences of countries b and c are not too strong, the threat of a bilateral FTA between countries b and c and the discrimination that is inherent to such trade agreement can be necessary to convince country a to announce in favor of the global free trade. When

26

such is the case, Article XXIV acts as a strong building bloc for achieving the global free trade.

Part (iii) of Proposition 6 provides the opposite argument to the second part: when countries are relatively symmetric in their political preferences which are moderately strong (or as we converge to the complete symmetry as in Proposition 3), the tariff level effect dominates the discrimination effect and the degree of positive externality enjoyed by country a is stronger under the multilateral agreement $\langle bc^m \rangle$ relative to the bilateral FTA $\langle bc \rangle$. Therefore, the incentives for participating in global free trade are stronger when discriminatory bilateral agreements are not permitted. In this case the Article XXIV acts as a strong stumbling bloc for the global free trade.

From the above discussion, we know that when $\beta_a > \max\{\beta_a(F - bc^m)_a, \beta_a(F - bc)_a\}$, global free trade fails to obtain under both bilateralism and multilateralism. Intuitively, when the the degree of political asymmetry is sufficiently large, then even the possibility of a bilateral FTA between the countries with weaker political preference is not enough to induce country a to opt for global free trade. In such case, it is informative to compare global welfare from the equilibrium agreements that obtain under bilateralism and multilateralism. From Propositions 4 and 5 we know that when $\max\{\beta_a(F - bc^m)_a, \beta_a(F - bc)_a\} \leq \beta_a$ and $\beta \leq \beta(bc - \Phi)_b$ hold, bilateral FTAs obtain in the equilibrium under bilateralism and an MA obtains under multilateralism. Since the degree of trade liberalization is larger under bilateral FTAs relative to two-country MFN-based agreements, global welfare is greater under the bilateralism game than in the multilateralism game. Thus, when free trade is out of reach, the option to pursue bilateral FTAs can yield welfare-improving trade liberalization that is foregone under a strictly multilateral approach. In this case Article XXIV acts as a partial building bloc.

On the other hand, when $\beta > \beta(bc - \Phi)_b$ holds, while two country MFN-based agreements arise in CPNE under the multilateralism game, No Agreement $\langle \Phi \rangle$ is a unique CPNE under the bilateralism game. Since partial trade liberalization is always better than no trade liberalization, we obtain higher welfare when countries follow only a multilateral approach to trade liberalization. Under this scenario, the existence of Article XXIV leads to a partial stumbling bloc.

6   Conclusions

During a period when multilateral trade liberalization seems to have come to a grinding halt, the question whether PTAs have a facilitating role to play in the multilateral trading system becomes even more important. PTAs appear to be the only type of reciprocal trade liberalization countries are interested in undertaking nowadays. While member countries reciprocally reduce internal tariffs on each other under preferential

trade liberalization, discriminatory treatment toward non-members is inherent to such trade agreements. As a result, the concerns that the pursuit of PTAs might undermine multilateral trade liberalization efforts remain. This paper provides a political economy analysis of countries' decisions to pursue preferential and multilateral trade liberalization, and analyzes the role of the GATT article XXIV, which sanctions PTAs as an exception to non-discrimination principle, for the prospects of global free trade. We develop a model with endogenous formation of trade agreements, both preferential and multilateral, where governments choose their trade policies optimally under any trade regime. In contrast with previous studies, we allow policymakers to have political motivations so that they value industry interests potentially more than social welfare. In line with the existing empirical evidence, we also allow for political preferences to vary across countries. By comparing equilibrium outcomes under bilateralism and multilateralism, we identify the effect of political forces and GATT article XXIV on the global trade system and on the likelihood of achieving the global free trade.

Our analysis reveals that the pressure applied by industrial interest groups and heterogeneity in preferences of policymakers across countries play an important role in judging the success of bilateralism and multilateralism. We show that when political motivations are not extreme (neither too strong nor too weak) and the variation in political preferences across countries is low, a bilateral FTA emerges as a unique equilibrium outcome, whereas global free trade would have prevail had preferential trade liberalization not been available. This strong stumbling block effect of FTAs owes to substantial reduction in tariffs by FTA member countries that motivates other countries to stay out of the agreement and free ride on trade liberalization effort by FTA members. However, with heterogeneous political preferences, the freedom to pursue preferential agreement may become a necessary condition for global free trade so that FTAs act as a strong building block to multilateral liberalization. This would be the case when the political-economy motives for protection are not too strong so that the external tariffs of countries forming a bilateral or a multilateral agreement are not substantially different. Then a third country would choose to join the FTA in order to avoid tariff discrimination within the agreement, as long as the political bias of a third country is not too large for it to opt out of the agreement entirely. Therefore, the threat of preferential trade liberalization and discrimination inherent to such trade liberalization may be necessary to induce countries with stronger special interest groups to participate in multilateral free trade negotiations.

Appendix A: Proofs

Here we report the key formulae that are necessary for proving our results. For an arbitrary tariff vector $\mathbf{t} = (t_{ij}, t_{ik}, t_{ji}, t_{jk}, t_{ki}, t_{kj})$, we can write country i's government payoff as

$$g_i = CS_i + (1+\beta_i)\pi_{ii} + \sum_{z=j,k} \pi_{iz} + TR_i,$$

where consumer surplus in country i equals

$$CS_i = \frac{1}{2}\left[\frac{3(\alpha - c) - t_{ij} - t_{ik}}{4}\right]^2$$

while domestic and foreign export profits are

$$\pi_{ii} = \left[\frac{(\alpha - c) + t_{ij} + t_{ik}}{4}\right]^2$$

$$\sum_{z=j,k} \pi_{iz} = \left[\frac{(\alpha - c) - 3t_{ji} + t_{jk}}{4}\right]^2 + \left[\frac{(\alpha - c) - 3t_{ki} + t_{kj}}{4}\right]^2$$

and the tariff revenue is given by

$$TR_i = \frac{t_{ij}[(\alpha - c) - 3t_{ij} + t_{ik}] + t_{ik}[(\alpha - c) - 3t_{ik} + t_{ij}]}{4}$$

Using the above formulae and the optimal tariff levels reported in the text, we can easily calculate government payoffs under all possible trade agreements. Note that, under a multilateral free trade, we replace $\beta_i = 0$, leading to the equivalence of the government's payoff and national welfare. To save space, we do not include the algebraic details underlying these straightforward calculations but they are available upon request.

Proof of Lemma 1

Assuming $\beta_i = \beta \leq \frac{1}{2}$, the following obtains:

Part *(i)*: Using the above welfare components and optimal tariffs under the given trade regimes, when $\beta_i = \beta \leq \frac{1}{2}$ holds, the following obtains for all β: $\Delta g_i(ih - F) > 0$; $\Delta g_i(ih - jh) > 0$; $\Delta g_i(ih - ij) > 0$; $\Delta g_i(ih - jk) > 0$ and $\Delta g_i(ih - \Phi) > 0$ while $\Delta g_k(ih - ij) < 0$.

Part *(ii)* $\Delta g_i(ij - \Phi) = \Delta g_j(ij - \Phi) \geq 0$ *iff* $\beta \leq \beta(ij - \Phi)_i \cong 0.423$;

Part *(iii)* $\Delta g_j(F - ih) = \Delta g_k(F - ih) \geq 0$ *iff* $\beta \leq \beta(F - ih)_j \cong 0.447$ *and*

Part *(iv)* $\Delta g_i(F - jk) \geq 0$ *iff* $\beta \leq \beta(F - jk)_i \cong 0.214$

Inequality 13: $\Delta g_k(ij - \Phi) = \frac{(\alpha - c)^2(51 - 76\beta + 12\beta^2)}{2(21 - 2\beta)^2(5 - 2\beta)^2} > 0$ *for all* $\beta \leq \frac{1}{2}$.

Inequality 14: $\Delta g_i(F - ij) \geq 0$ *iff* $\beta \leq \beta(F - ij)_i \cong 0.274$.

Inequality 15: $\Delta g_i(F - \Phi) \geq 0$ *iff* $\beta \leq \beta(F - \Phi)_i \cong 0.356$.

Proof of Proposition 1

First note from the proof of Lemma 1 that $\beta(ij - \Phi)_i \cong 0.423 > \beta(F - jk)_i \cong 0.214$. Given that $\beta \geq \beta(ij - \Phi)_i \cong 0.423$, we know from inequality (15) that there exists no deviation from No Agreement $\langle \Phi \rangle$ to free trade. Finally, $\Delta g_j(ih - \Phi) \geq 0$ *iff* $\beta \leq \beta(F - ij)_i \cong 0.3$ and thus when $\beta \geq \beta(ij - \Phi)_i \cong 0.423$, there exists no deviation from No Agreement $\langle \Phi \rangle$ to any hub and spoke regime.

Inequality 19: $g_k(ij) \leq g_k(ij^m)$ when $\beta \leq \beta(ij - ij^m)_k \cong 0.146$

Inequality 20: $\Delta g_i(ij^m - \Phi) = \Delta g_j(ij^m - \Phi) = \frac{(\alpha - c)^2(1 - 2\beta)^2}{2(3 - 2\beta)(5 - 2\beta)^2} \geq 0$ *for all* $\beta \leq \frac{1}{2}$.

Inequality 21: $\Delta g_k(F - ij^m) > 0$ *iff* $\beta < \beta(F - ij^m)_k \cong 0.303$

Proof of Proposition 3

It is immediate from the proof of Lemma 1 and inequality 21 that $\beta(ij - \Phi)_i \cong 0.423 > \beta(F - ij^m)_k \cong 0.303 > \beta(F - ij)_k \cong 0.214$. The rest of the proposition is immediate from Propositions 1 and 2.

Proof of Lemma 2: Given that $\beta_i, \beta_j \leq \frac{1}{2}$ holds, using the above welfare components and optimal tariffs under the given trade regimes, the following obtains:

$$
\begin{aligned}
\frac{\partial \Delta g_i(ij - \Phi)}{\partial \beta_i} &= \frac{\partial \Delta g_i(jh - jk)}{\partial \beta_i} < 0, \quad \frac{\partial \Delta g_i(ij - \Phi)}{\partial \beta_j} = \frac{\partial \Delta g_i(ih - ik)}{\partial \beta_j} > 0, \\
\frac{\partial \Delta g_i(ih - ik)}{\partial \beta_i} &= \frac{\partial \Delta g_i(F - kh)}{\partial \beta_i} < 0 \text{ and } \frac{\partial \Delta g_i(jh - jk)}{\partial \beta_j} = \frac{\partial \Delta g_i(F - kh)}{\partial \beta_j} > 0
\end{aligned}
$$

Proof of Lemma 3: Given that $\beta_i, \beta_j \leq \frac{1}{2}$ holds, using the above welfare components and optimal tariffs under the given trade regimes, the following obtains:

$$
\frac{\partial \Delta g_k(ih - ik)}{\partial \beta_i} = \frac{\partial \Delta g_k(F - kh)}{\partial \beta_i} < 0
$$

and

$$
\frac{\partial \Delta g_k(ih - ik)}{\partial \beta_j} = \frac{\partial \Delta g_k(ij - \Phi)}{\partial \beta_i} > 0 \text{ iff } \beta_j \gtrsim \overline{\beta}
$$

and the critical β in part (ii) is found as $\overline{\beta} = 0.271$.

Proof of Lemma 4: Given that $\beta_i, \beta_j, \beta_k \leq \frac{1}{2}$ holds, using the above welfare

components and optimal tariffs under the given trade regimes, the following obtains:

$$\frac{\partial \Delta g_i(ij^m - \Phi)}{\partial \beta_i} \geq 0, \ \frac{\partial \Delta g_i(ij^m - \Phi)}{\partial \beta_j} \leq 0, \ \frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_i} < 0,$$

$$\frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_j} \geq 0 \text{ and } \frac{\partial \Delta g_i(F - ij^m)}{\partial \beta_k} \geq 0$$

Critical Political Biases and Proofs of Propositions 4, 5 and 6

Using the welfare levels reported above, we can easily find the critical political bias values that are also represented in the figures in the text. First we focus on the CPNE condition for $\langle F \rangle$ and find that country a has an incentive to unilaterally deviate from its announcement $\{b, c\}$ to $\{\phi, \phi\}$ that leads to a deviation from $\langle F \rangle$ to $\langle bc \rangle$ when $\beta_a > \beta_a(F - bc)_a$ holds and it is a self enforcing deviation. Moreover, using the welfare levels reported above, we can easily show that there exist no other unilateral or coalitional announcement deviations from $\langle F \rangle$ and thus free trade $\langle F \rangle$ is a CPNE when $\beta_a \leq \beta_a(F - bc)_a$.

Next, we show that the joint deviation of countries b and c from their respective announcements $\{\phi, \phi\}$ and $\{\phi, \phi\}$ to $\{\phi, c\}$ and $\{\phi, b\}$, leading to a deviation from $\langle \Phi \rangle$ to $\langle bc \rangle$, happens when $\beta < \beta(bc - \Phi)_b \cong 0.423$ and it is self enforcing. Using the welfare levels reported above, when $\beta \geq \beta(bc - \Phi)_b$ holds, there exists no other coalitional deviations from the Nash equilibrium strategy profile leading to No Agreement $\langle \Phi \rangle$, and thus $\langle \Phi \rangle$ is a CPNE. Now consider whether bilateral FTAs that involve country a (i.e. $\langle ab \rangle$ and $\langle ac \rangle$) are CPNE. First note that coalitional announcement deviations from $\langle ab \rangle$ (or $\langle ac \rangle$) to $\langle F \rangle$ is self-enforcing only when $\beta_a \leq \beta_a(F - bc)_a$. Second, country a has an incentive to unilaterally deviate from its announcement $\{b, \phi\}$ ($\{\phi, c\}$) to $\{\phi, \phi\}$ that leads to a deviation from $\langle ab \rangle$ (or $\langle ac \rangle$) to $\langle \Phi \rangle$ when $\beta_a < \beta_a(ab - \Phi)_a$. Finally, the coalitional announcement deviation of countries a and c from $\langle ab \rangle$ to $\langle ah \rangle$ occurs when $\beta_a > \beta_a(ah - ab)_c$ holds and it is a self-enforcing deviation.

Combining the above critical political biases, we obtain that $\langle ab \rangle$ and $\langle ac \rangle$ are CPNE when $\min\{\beta_a(ah - ab)_c, \beta_a(ab - \Phi)_a\} \geq \beta_a \geq \beta_a(F - bc)_a$ holds. We can also show over the same range that $\langle bc \rangle$ is a CPNE as well (while not unique). Finally, using these critical political biases, we can show that there exists no unilateral and coalitional deviation from $\langle bc \rangle$ when (a) $\beta_a > \beta(F - bc)_a$ and (b) $\beta_a > \min\{\beta_a(ah - ab)_c, \beta_a(ab - \Phi)_a\}\}$ and (c) $\beta_a < \beta_a(bc - \Phi)_b$ and thus $\langle bc \rangle$ is the unique CPNE.

Similarly, in the multilateralism game, country a has an incentive to unilaterally deviate from its announcement of $\{M\}$ to the announcement $\{\phi\}$ (leading to a deviation from $\langle F \rangle$ to $\langle bc^m \rangle$) when $\beta_a > \beta_a(F - bc^m)_a$ holds and there exist no other unilateral or coalitional announcement deviations from $\langle F \rangle$ and thus free trade $\langle F \rangle$ is a CPNE when $\beta_a \leq \beta_a(F - bc^m)_a$ holds.

Note from the inequality (20) and Lemma 4 that countries b and c have incentives to jointly deviate from their announcements of no agreement to announcements of multilateralism and this deviation from $\langle \Phi \rangle$ to $\langle bc^m \rangle$ is self-enforcing. Thus, No Agreement $\langle \Phi \rangle$ never arises as a CPNE in the multilateralism game. Now consider whether two-country multilateral agreement that involve country a (i.e. $\langle ab^m \rangle$ and $\langle ac^m \rangle$) are CPNE. Focusing on $\langle ab^m \rangle$ we have that country a has no incentive to unilaterally deviate while country b has an incentive to unilaterally deviate from its announcement of multilateralism to no agreement when $\beta_a > \beta_a(ab^m - \Phi)_b$. Moreover, when $\beta_a > \beta_a(F - ab^m)_c$ holds, country c has an incentive to unilaterally deviate from its announcement of no agreement to announcement of multilateralism leading to a deviation from $\langle ab^m \rangle$ to $\langle F \rangle$ and this deviation is self-enforcing.

Then, we argue that $\langle ab^m \rangle$ and $\langle ac^m \rangle$ are CPNE when $\beta_a \leq \min\{\beta_a(ab^m - \Phi)_b, \beta_a(F - ab^m)_c\}$ holds. We can also show over the same range that $\langle bc^m \rangle$ is a CPNE as well (while not unique). Finally, using these critical political biases, we can show that there exists no unilateral and coalitional deviation from $\langle bc^m \rangle$ when (a) $\beta_a > \beta_a(F - bc^m)_a$ and (b) $\beta_a > \min\{\beta_a(ab^m - \Phi)_b, \beta_a(F - ab^m)_c\}$ hold and thus $\langle bc^m \rangle$ is the unique CPNE.

The proof of Proposition 6 is immediate from the comparisons of Proposition 4 and Proposition 5.

Appendix B: Alternative Pattern of Asymmetry

In the main text we modelled the cross-country differences in political preferences in a way that one country has stronger political bias than the other two. In this appendix we show that the main result of the paper about strong building and strong stumbling block effects of FTAs survives for an alternative definition of asymmetry in political preferences when majority of the countries have strong political biases. Specifically, here we assume that the pattern of asymmetry in political motivations is given by: $\beta_a \leq \beta_b = \beta_c = \beta \leq \frac{1}{2}$. To avoid redundancy, we directly state our main result that is a modified version of our Proposition 6, illustrated in Figure 5:

Proposition 7: *Suppose $\beta_a \leq \beta_b = \beta_c = \beta \leq \frac{1}{2}$. Then, the following obtains:*

(i) (Irrelevance of Bilateralism) when $\beta_a > \max\{\beta_a(F - ab^m)_c, \beta_a(F - ab)_c\}$ holds, $\langle F \rangle$ is a CPNE under both bilateralism and multilateralism.

(ii) (Strong Building Block) when $\beta_a(F - ab)_c < \beta_a < \beta_a(F - ab^m)_c$ holds, $\langle F \rangle$ is a CPNE is under bilateralism but not under multilateralism;

(iii) (Strong Stumbling Block) when $\beta_a(F - ab^m)_c < \beta_a < \beta_a(F - ab)_c$ holds, $\langle F \rangle$ is a CPNE is under multilateralism but not under bilateralism.

– Figure 5 –

32

It is important to note from the comparison of Propositions 6 and 7 that, irrespective of the nature of asymmetry in political preferences, similar results obtain regarding the relative merits of bilateralism and multilateralism for achieving global free trade. As in Proposition 6, the first part of the above proposition implies that when both the degree of political preferences of countries and the degree of asymmetry in political preferences are sufficiently small, free trade arises as a CPNE under both regimes and thus the freedom to pursue bilateral FTAs plays no potential role for the ultimate objective of achieving global free trade.

Similar to our previous discussion in Proposition 6, whether free trade is a CPNE or not depends critically on the unilateral deviation incentive of the countries with stronger political preference (say country c). Therefore, we still focus on the comparison of country c's government payoff under $\langle ab^m \rangle$ and $\langle ab \rangle$. As before, while country c faces higher tariff under $\langle ab^m \rangle$ than under $\langle ab \rangle$, it is discriminated against its rival exporter in each member country's market under the bilateral FTA $\langle ab \rangle$ but not under $\langle ab^m \rangle$. Note from the second part of the proposition that FTAs act as a strong building bloc when the asymmetry in political preferences is sufficiently large and the political preference of country a is not too strong. The intuition behind this result mimics our discussion in Proposition 6: over this parameter range the discrimination effect dominates the tariff level effect and thus country c has less incentive to unilaterally deviate from its free trade announcement under bilateralism relative to multilateralism.

References

[1] Aghion, Philippe, Pol Antràs and Elhanan Helpman, 2007. "Negotiating Free Trade." *Journal of International Economics* 73, 1-30.

[2] Bagwell, Kyle and Robert. W. Staiger, 1997. "Multilateral Tariff Cooperation During the Formation of Free Trade Areas." *International Economic Review* 38, 291-319.

[3] Bagwell, Kyle and Robert. W. Staiger, 1999. *Regionalism and multilateral tariff cooperation.* In John Pigott and Alan Woodland, eds., International Trade Policy and the Pacific Rim, London: Macmillan.

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

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

[6] Bhagwati, Jagdish, Arvind Panagariya, and Pravin Krishna, eds., Trading Blocs, 1999, The MIT Press, Cambridge, MA.

[7] Bond, Eric W., Raymond G. Riezman, and Constantinos Syropoulos, 2004. "A Strategic and Welfare Theoretic Analysis of Free Trade Areas." *Journal of International Economics* 64, 1-27.

[8] Cadot, Olivier, Grether, Jean-Marie, and Marcelo Olarreaga, 2003. "India's Trade Policy for Sale: How Much? Who Buys?" CEPR discussion paper 4168.

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

[10] Estevadeordal, Antoni, Caroline Freund and Emanuel Ornelas, 2008. "Does Regionalism Affect Trade Liberalization toward Non Members?" *Quarterly Journal of Economics* 123(4), 1531-1575.

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

[12] Gawande, Kishore, and Usree Bandyopadhyay, 2000. "Is Protection for Sale? Evidence on the Grossman-Helpman Theory of Endogenous Protection." *Review of Economics and Statistics* 82, 139–52.

[13] Gawande, Kishore, Pravin Krishna, and Marcelo Olarreaga, 2009. "What Governments Maximize and Why: The View from Trade." *International Organization* 63, 491-532.

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

[15] Grossman, Gene. M., Elhanan Helpman, 1994. "Protection for Sale." *American Economic Review* 84(4), 833–850.

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

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

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

[19] Krugman, Paul R. "The Move Toward Free Trade Zones." in Policy Implications of Trade and Currency Zones: A Symposium Sponsored by the Federal Reserve Bank of Kansas City, Federal Reserve Bank of Kansas City, Kansas City, 7-41, 1991.

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

[21] McCalman, Phillip, 2004. "Protection for Sale and Trade Liberalization: An Empirical Investigation." *Review of International Economics* 12, 81–94

[22] Mitra, Devashish, Dimitrios Thomakos, and Mehmet Ulubasoglu, 2002. "Protection for Sale in a Developing Country: Democracy versus Dictatorship." *Review of Economics and Statistics* 84, 497–508.

[23] Ornelas, Emanuel, 2005a. "Trade Creating Free Trade Areas and the Undermining of Multilateralism." *European Economic Review*, 49(7), pages 1717-35.

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

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

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

[27] Saggi, Kamal and Halis M. Yildiz, 2010. "Bilateralism, Multilateralism, and the Quest for Global Free Trade." *Journal of International Economics* 81, 26-37.

[28] Saggi, Kamal and Halis M. Yildiz, 2011. "Bilateral Trade Agreements and the Feasibility of Multilateral Free Trade." *Review of International Economics*, 19, 356-373.

[29] Saggi, Kamal, Alan Woodland and Halis M. Yildiz, 2013. "On the Relationship between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions." *American Economic Journal: Microeconomics*, 5(1), 63-99.

[30] Seidmann, Daniel J., 2009. "Preferential Trading Arrangements as Strategic Positioning." *Journal of International Economics* 79, 143–159.

[31] Stoyanov, Andrey, 2009. "Trade Policy of a Free Trade Agreement in the Presence of Foreign Lobbying." *Journal of International Economics* 77, 37-49.

[32] Stoyanov, Andrey, 2014. "Endogenous Trade Agreements and Foreign Lobbying." *Review of International Economics* 22, 561-577.
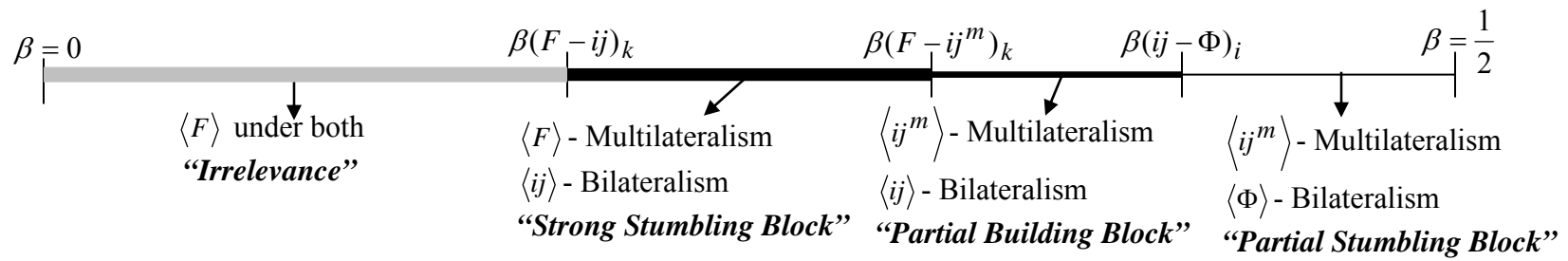
Figure 1: Bilateralism vs. Multilateralism under Symmetry
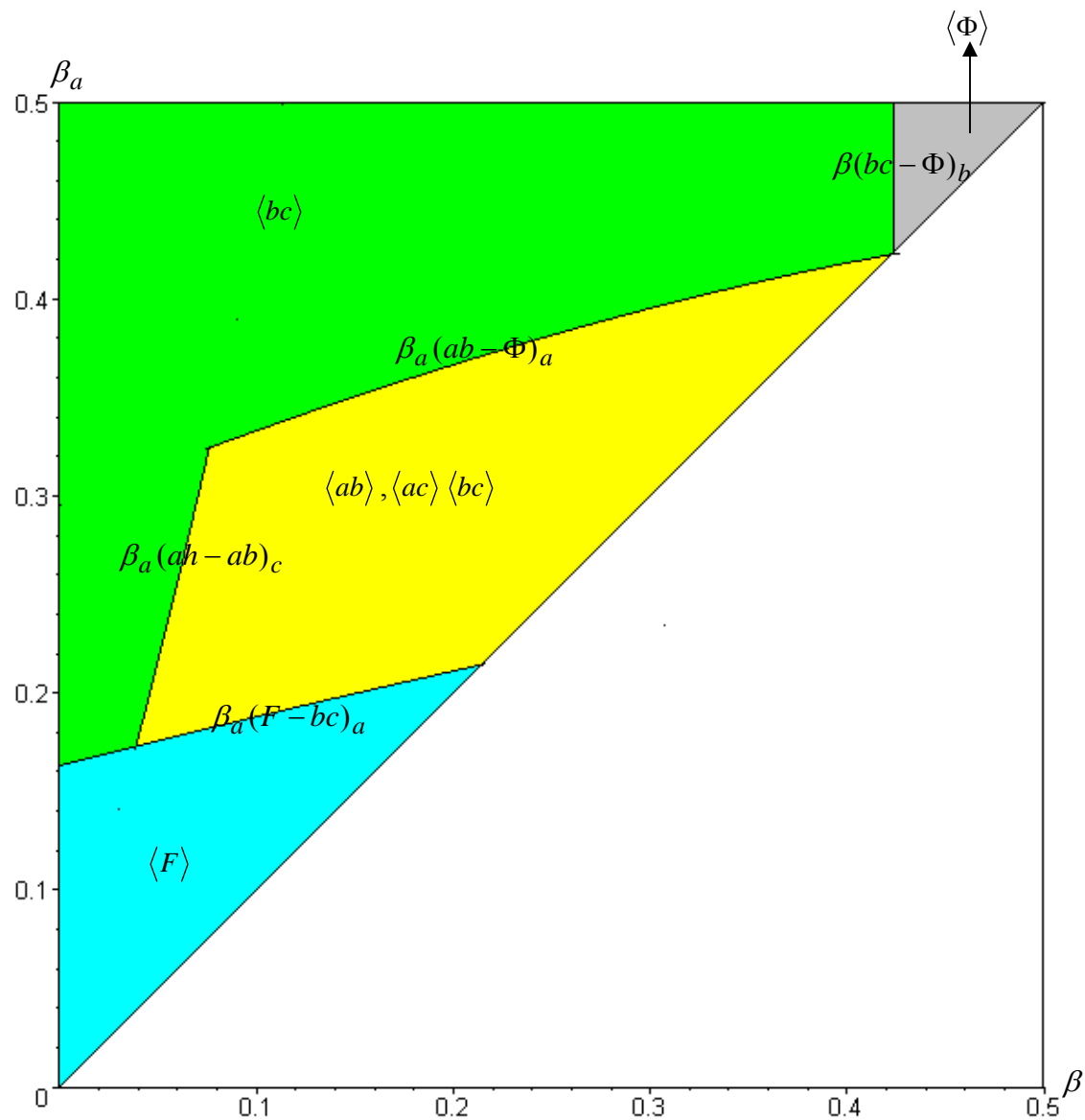
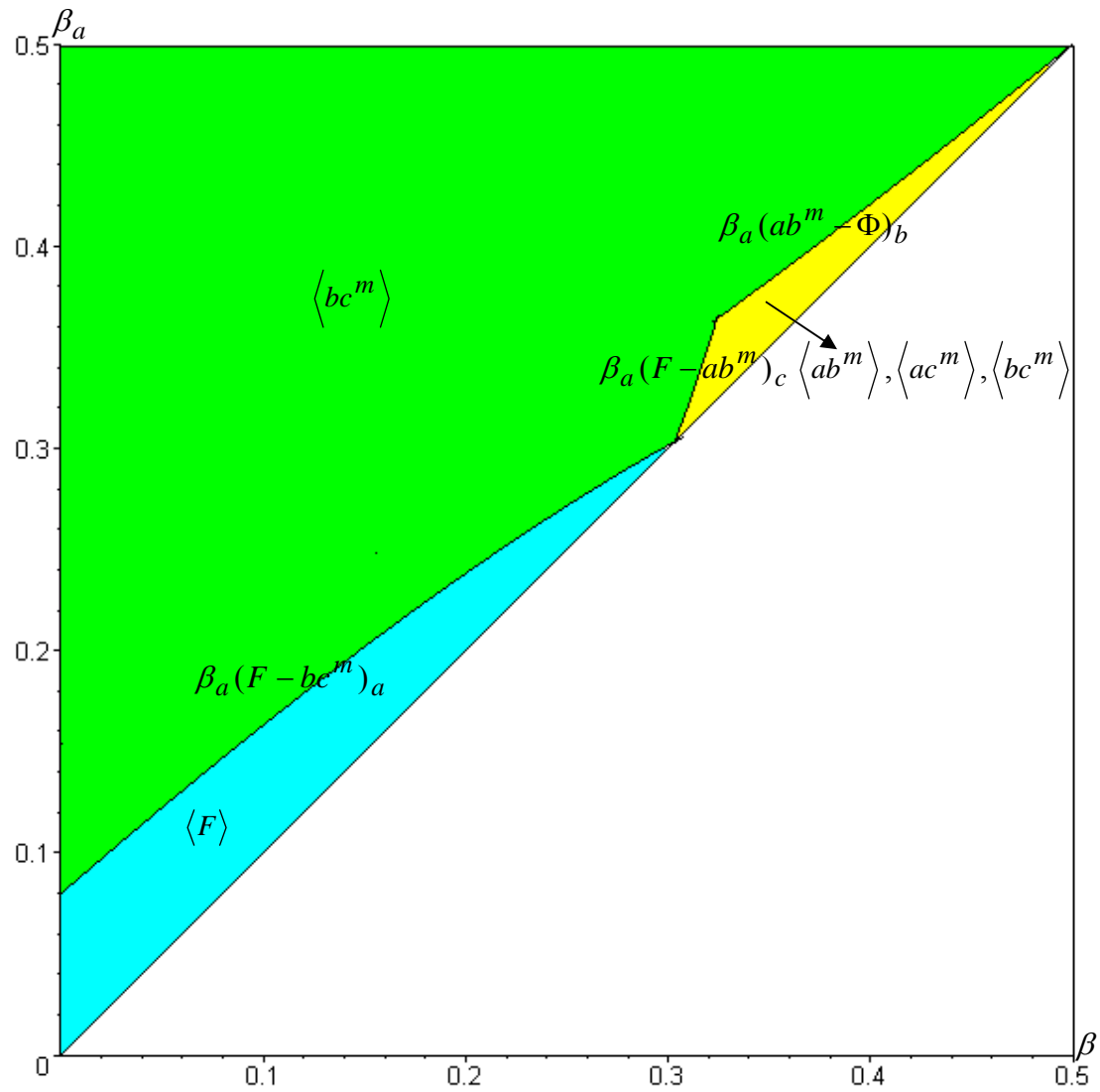Figure 2: CPNE Agreements under Bilateralism – Asymmetry

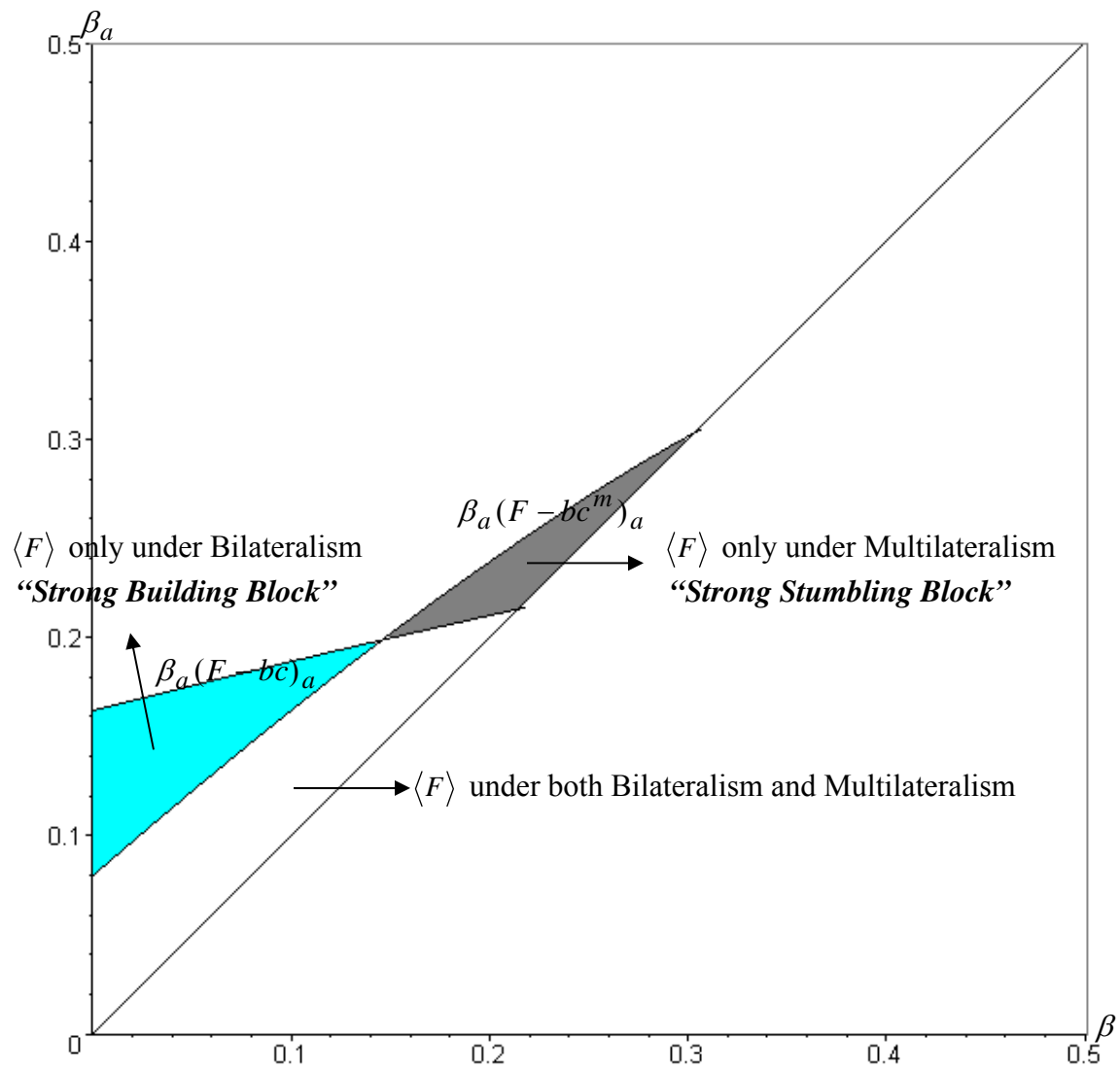Figure 3: CPNE Agreements under Multilateralism – Asymmetry

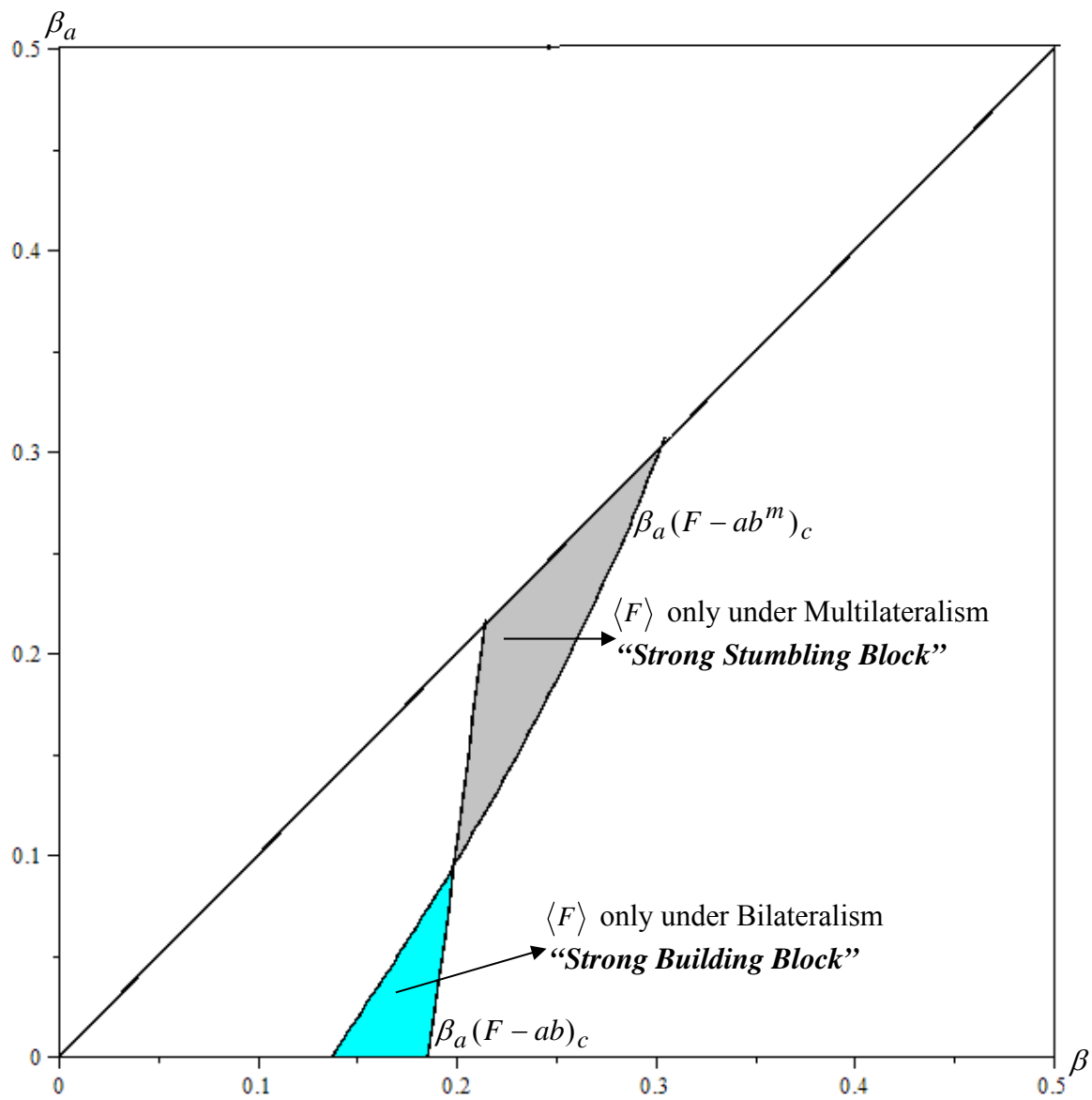Figure 4: Bilateralism versus Multilateralism under Asymmetry

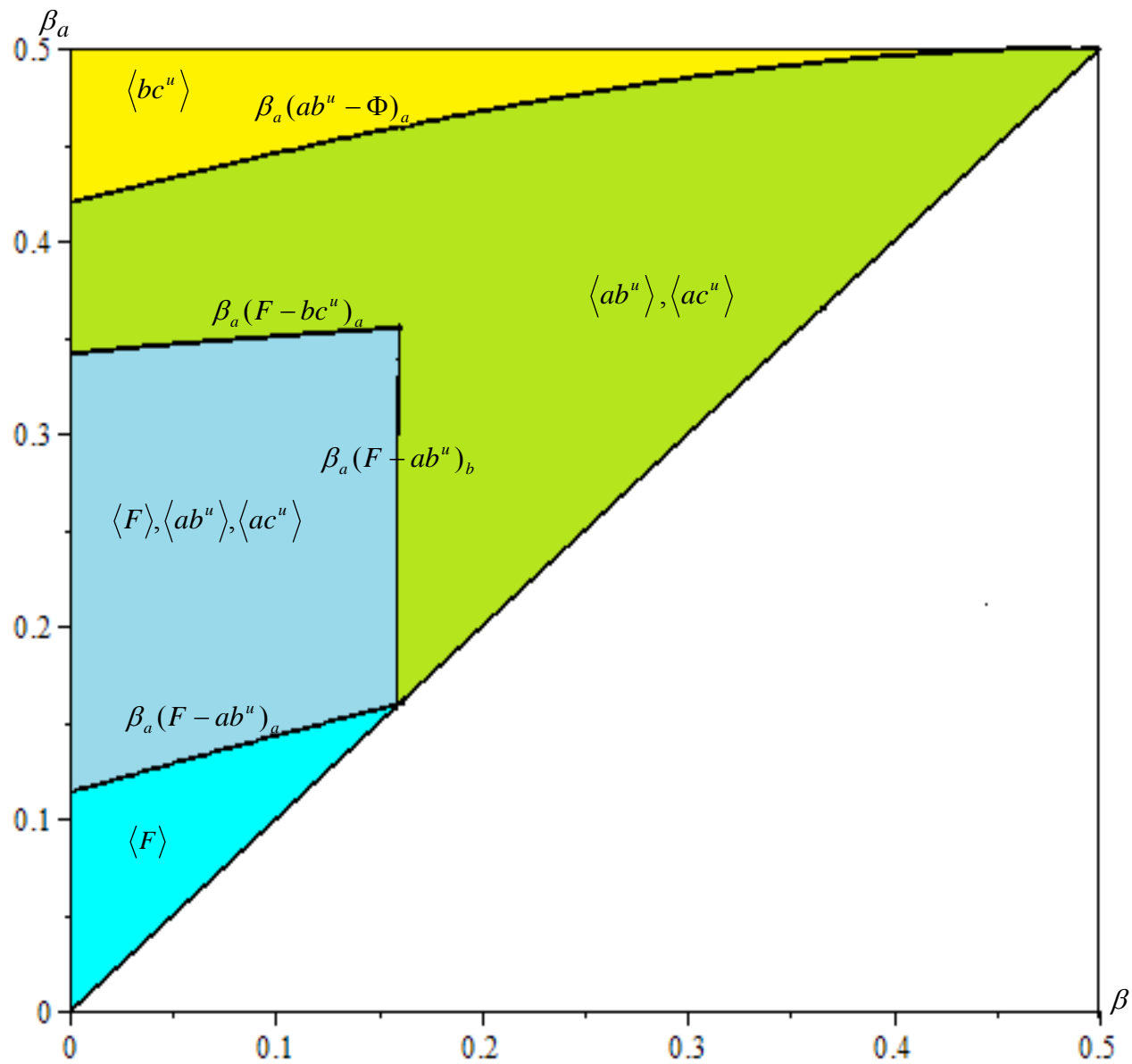Figure 5: Bilateralism versus Multilateralism under Alternative Pattern of Asymmetry
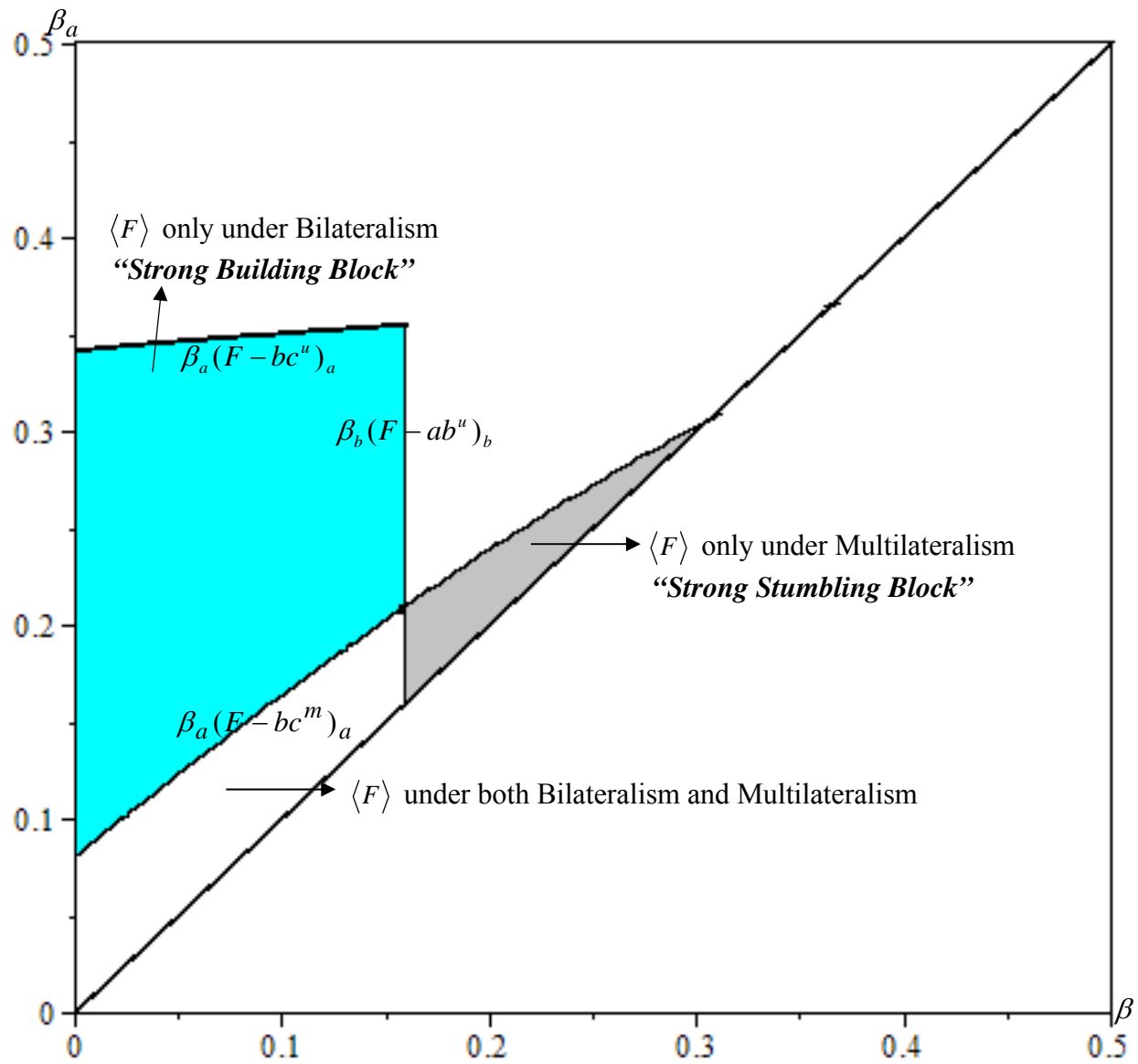
Figure 6: CPNE Agreements under CU Game − Asymmetry

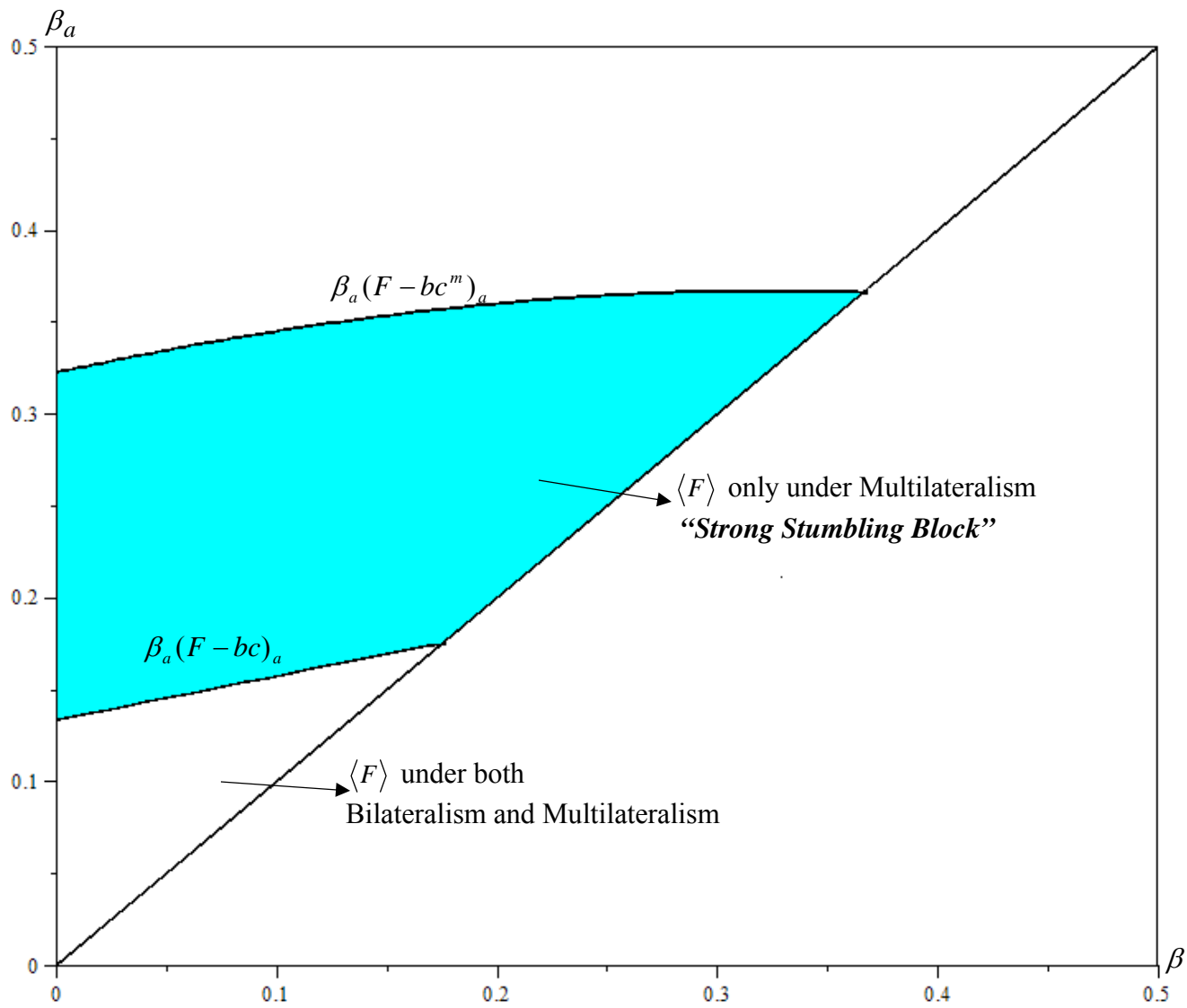Figure 7: Bilateralism versus Multilateralism under CU Game – Asymmetry

Figure 8: Bilateralism versus Multilateralism with Lobbying for Trade Regime