Preferential trade agreements as building or stumbling blocs?
The importance of commitment

James Lake*

Johns Hopkins University

February 2011

Abstract

Will continual expansion of preferential trade agreements (PTAs) lead to global free trade? In the presence of multilateral negotiations, are PTAs necessary for, or will they necessarily lead away from, global free trade? This paper shows both answers depend crucially on a commitment problem faced by PTA members. Because the commitment problem revolves around a desire to commit to *not* form future PTAs, global free trade is not attained when PTA member's overcome their commitment problem. Market size asymmetry is an important determinant of member's ability to overcome the commitment problem. Higher levels of asymmetry generally increase member's rents and thus their ability to overcome it. This reduces the scope for global free trade. However, there is a threshold level of asymmetry at which the commitment problem cannot be overcome. This dramatically increases the extent to which PTAs both lead to, and are necessary for, global free trade.

JEL: C71, F12, F13

Keywords: Bilateral trade agreements, multilateralism, free trade, networks, commitment problem, farsighted

1 Introduction

Following the decision of the U.S. in the mid 1980's to pursue trade liberalization through preferential trade agreements (PTAs), the number of such agreements has continued to expand at a significant rate. According to the WTO website, the WTO had been notified of over 370 PTAs as of July 2010 and "the surge in [PTAs] has continued unabated since the early 1990s".[1] This surge quickly led economists such as Bhagwati (1991,

*Email: jlake2@jhu.edu. I would like to especially thank Pravin Krishna and M. Ali Khan for very useful guidance, advice and discussion. I would also like to thank Kamal Saggi and Matt Jackson for useful comments and discussion.

[1] http://www.wto.int/english/tratop_e/region_e/region_e.htm, October 7, 2010.

1993) to argue that PTAs are "stumbling blocs" because they create incentives which hinder the movement towards global free trade. Other economists such as Summers (1991) and Baldwin (1996) argued PTAs promote further liberalization and are thus "building blocs".

Initially, the building bloc–stumbling bloc terminology came to be understood as a question of how an exogenously given bilateral agreement (BA), which makes members myopically better off, affects the incentives for further multilateral liberalization. Levy (1997) shows that if the public is fairly indifferent between the status quo and global free trade, such a BA can hinder multilateralism if it leads to disproportionate gains for one group. Krishna (1998) shows BA members may block global free trade due to the rent dissipation incurred in moving to global free trade even though they would not have blocked this move in the absence of BAs. In contrast, Ornelas (2005) shows the presence of tariff complementarity can mean BAs mitigate the political incentives that prohibit global free trade and thus strengthen multilateral incentives. In addition to these static models, models of repeated interaction which emphasize the importance of self enforcing multilateral tariff cooperation have also been used to analyze this question — e.g. Bagwell and Staiger (1997), Fruend (2000), Bond et. al. (2001) and Saggi (2006).

This paper contributes to a more recent strand of the literature in which the building bloc–stumbling bloc question has come to be understood in the context of equilibrium determination of both the degree and nature of trade liberalization (Riezman (1999), Aghion et. al. (2007), Saggi and Yildiz (2010, 2011), Saggi et. al. (2010)). In this literature, as well as this paper, the nature of liberalization takes one of two forms. The first is the most favored nation (MFN) principle of GATT Article I; any tariff reduction afforded to one country must be afforded to all. The second is the "escape clause" governed by GATT Article XXIV; reciprocal elimination of tariffs between PTA members is permitted if barriers to trade on third party countries are not increased. Comparison of the equilibrium when both the MFN principle and PTAs are avilable with the equilibrium when only the MFN principle is available reveals whether PTAs are "strong building blocs", because they are necessary to attain global free trade, or whether they are "strong stumbling blocs", because global free trade is only attained in their absence.[2] The current paper also adopts this comparison approach.

To investigate the strong building bloc–strong stumbling bloc question, this paper draws on network theory to present a novel dynamic network model in which countries are farsighted. In doing so, it develops a new equilibrium concept; a farsighted dynamic network equilibrium. The three country model is dynamic because countries have the option to form BAs or MFN consistent agreements sequentially, at most one per

[2] Until now in the introduction, the precise meaning of PTAs being building blocs and stumbling blocs has been left somewhat vague. This reflects the fact that, throughout the literature, there has been variation in the definition attached to this terminology (e.g. Saggi (2006), Aghion et al. (2007) and Furusawa and Konishi (2007)). The usage of "strong stumbling bloc" and "strong building bloc" is due to Saggi and Yildiz (2011) in recognition of this fact.

period, and a country's overall payoff is a discounted sequence of one period payoffs based on oligopolistic competition. In addition, countries are farsighted because they anticipate the equilibrium evolution of the global trade network conditional upon forming and not forming a potential trade agreement under consideration, and compare the two discounted payoffs. To determine which countries form a trade agreement in any given period, the model employs a coalitional simultaneous move game utilizing the equilibrium concept of Page et. al. (2005). By nesting the coalitional simultaneous move game within an overall dynamic network framework, where this dynamic network framework is motivated by Dutta et. al. (2005), the new equilibrium concept emerges.

The farsighted and dynamic aspects of the equilibrium concept shed new light on the strong building bloc–strong stumbling bloc question. Because of these aspects, the importance of a commitment problem arises. This commitment problem has not been identified in the literature and relates to the desire of two countries who have the sole BA, referred to as insiders, to commit to *not* form another BA. If an insider forms the second BA, it becomes the "hub" while the other countries become "spokes". The rents associated with this additional preferential access imply being a hub is preferable to being an insider. However, if one insider will become the hub, then the other insider will also want to become the hub to avoid becoming a spoke. Yet, if the two spokes form their own BA which leads to global free trade, then, depending on the discount factor, *both* insiders can be worse off. In this situation, each insider has an incentive to commit to not become the hub. Largely speaking, BAs lead to global free trade if and only if the two largest countries, i.e. those with the largest market size and greatest preferential access to offer, cannot overcome their commitment problem.

Market size asymmetry is a crucial determinant of insiders' ability to overcome their commitment problem. A striking result is that, essentially, while increasing degrees of asymmetry increase the ability of insiders to overcome their commitment problem through increasing the level of protected rents, there is a critical level of asymmetry where this ability experiences a dramatic once off drop. In turn, the extent to which BAs lead to global free trade rises dramatically. The threshold level of asymmetry is such that, as a spoke, the largest country will no longer form the final BA with the smallest country. This level of asymmetry is important because it means the medium sized country will not hesitate in becoming the hub because it no longer faces any consequences. This induces the largest country to become the hub which in turn leads to global free trade. The implication is that while greater levels of asymmetry initially decrease (increase) the extent to which BAs are strong building (stumbling) blocs, crossing this threshold level of asymmetry provides a dramatic reversal.

Even though this intuition has said nothing about MFN agreements or the comparison of equilibria, two insights from the model imply the commitment problem is the crucial element that drives the strong building bloc–strong stumbling bloc issue. First, because the model allows countries to form BAs sequentially, the rents earned along the equilibrium path mean countries choose BAs over MFN agreements. Second, when BAs do not exist, greater overall degrees of asymmetry merely increase the incentive for the largest country, who has the greatest preferential access to offer and the least to gain, to block the MFN agreement leading to global free trade. Thus, understanding how the equilibrium when only MFN agreements exist compares to the equilibrium when BAs also exist depends critically on the commitment problem that arises when BAs are the only form of liberalization.

From an international trade perspective, the two papers most closely related to this paper are Aghion et. al. (2007) and Saggi and Yildiz (2010).[3] Aghion et. al. (2007) develop a model where an exogenously given leader country dictates between sequential and multilateral liberalization and makes take it or leave it offers to other countries. They show that the leader's choice, and hence the answer to whether BAs are strong building blocs or strong stumbling blocs, depends on both the nature of coalition externalities, positive or negative, and whether global free trade maximizes the aggregate payoff. However, while the commitment problem is at the heart of this answer in the current paper, it is irrelevant for Aghion et. al. (2007) because the leader dictates trade agreement formation.

While Saggi and Yildiz (2010) allow any country to propose trade agreement formation with any other countries, their answer to the strong building bloc–strong stumbling bloc question still does not depend on the commitment problem. This is due to their solution concept of a coalition proof Nash equilibrium which implies countries are only "somewhat" farsighted. The commitment problem rests on the possibility, from the perspective of a deviating insider, of the other insider subsequently forming the final BA. Yet, this subsequent deviation is irrelevant for a coalition proof Nash equilibrium.[4] Another difference with Saggi and Yildiz (2010) is that while BAs can be strong building blocs they cannot be strong stumbling blocs in their model, whereas BAs can be either in this paper depending on whether insiders can overcome the commitment problem. The result of Saggi and Yildiz (2010) arises because when BAs exist in addition to MFN agreements, the effects of tariff complementarity imply the net positive coalition externality for a nonmember, and hence its outside option, is smaller than when only MFN agreements exist.

Because of its cooperative solution concept, this paper also shares similarities with other papers that have used cooperative solution concepts to analyze the effects of PTAs on the global trade network. Using numerical simulations, Riezman (1999) and Melatos and Woodland (2007) employ the core solution concept. While

[3] Mukunoki and Tachi (2006) show the possibility of a hub–spoke network and the incentive for a subsequent spoke–spoke BA, which leads to global free trade, increases the extent to which global free trade is attainable. However, because of the absence of coalition formation, commitment problems do not arise in their model. Moreover, because they do not model multilateralism, they do not analyze the strong building bloc–strong stumbling bloc issue. Finally, because their model is symmetric, they do not describe how asymmetry affects the extent to which global free trade arises.

[4] A coalition proof Nash equilibrium rules out a strategy profile as an equilibrium if there is a coalitional deviation which makes each member of the coalition better off and no sub coalition can then deviate and make each member better off. Because of the restriction to sub coalitions, which is acknowledged by Bernheim et. al. (1987, p.7), the concept is "somewhat" farsighted.

Riezman (1999) does compare the equilibrium of two games as this paper does, and both papers emphasize the importance of endowment asymmetries on the equilibrium of the global trade network, identification of the commitment problem depends on analytical solutions. Goyal and Joshi (2006) and Furusawa and Konishi (2007) are the first in the literature to use network theory. Unlike the dynamic network formation model of this paper, they use the static network theory framework of Jackson and Wolinsky (2005) from which emerges an emphasis on the positive role that concenssion diversion plays in allowing BAs to lead to global free trade. However, because BAs are the only form of trade liberalization in their models, they cannot address the strong building bloc–strong stumbling bloc issue.[5]

The rest of the paper proceeds as follows. Section 2 develops the network terminology and equilibrium concepts. Section 3 presents the underlying trade model. Section 4 analyzes the equilibrium when BAs are the only form of liberalization. Section 5 adds the possibility of MFN agreemnets and compares the resulting equilibrium with that when only MFN agreements exist to explore the strong building bloc–strong stumbling bloc question. Finally, Section 6 concludes.

2 The network formation games and equilibrium concepts

2.1 Overview

Three games are considered in this paper. As discussed in the introduction, one of the games only allows MFN agreements while another game also allows BAs. Per Saggi and Yildiz (2010), the former is called the *multilateralism game* and the latter is called the *bilateralism game*. While comparing the equilibria of these two games answers whether BAs are strong building blocs or strong stumbling blocs, the answer depends crucially on the commitment problem and the intuition underlying the commitment problem is easily brought when allowing BAs are the only form of liberalization. The game when only BAs are allowed is called the *pure bilateralism game.* Each of the three games is a farsighted dynamic infinite horizon game played between three countries, $\mathcal{N} = \{s, m, l\}$. Figure 1 (see Appendix C) depicts the set of possible networks in these games and the subscripts that are used later when referring to specific network positions. An edge between two countries, generically labelled i, j and k, indicates they have a trade agreement while absence of an edge indicates absence of an agreement. Several useful features of the model are worth noting.

The first feature is that the status quo remains forever once either global free trade is reached or no BA forms in a given period, which this happens in at most three periods because there are only three possible

[5] Concession diversion describes the fact that the value of preferential access in a partner member's market falls as the partner member enters more BAs. This term is due to Ethier (2004) who argues it undermines further liberalization. In contrast, Goyal and Joshi (2006) and Furusawa and Konishi (2007) exploit the fact that concession diversion lowers the cost of preferential access for the member entering additional BAs.

BAs. This follows from assuming that BAs are binding (to be discussed below), and the implication of the model that it is not optimal to delay formation of a trade agreement by "waiting". Because of this status quo property, analytical tractability of the model is maintained.

The second feature is that the model is not sensitive to an assumption regarding the order in which countries are allowed to form BAs. While this is a typical problem of sequential move games, the model can essentially be seen as one that takes the Dutta et. al. (2005) dynamic network formation game and embeds a coalitional simultaneous move game in each period to endogenize the order of negotiation.[6] The specific simultaneous move equilibrium concept used is the Page et. al. (2005) notion of a consistent set. In this concept, all feasible coalitional deviations from a network in the consistent set are deterred because they lead to a stable network in which the coalition is worse off. Deterrence can arise for two reasons. First, the initial coalitional deviation leads to a network in the consistent set in which the coalition is worse off. Second, a subsequent sequence of coalitional deviations leads to a network in the consistent set in which the initial deviating coalition is worse off. It should be pointed out that, despite relating to a simultaneous move framework, the notion of a consistent set also embodies farsightedness. Thus, in addition to the farsightedness associated with the dynamic dimension of the game, the equilibrium concept built here also embodies farsightedness its simultaneous dimensions.

However, the degree of farsightedness used in the model does not come without a trade off. Along the simultaneous dimension, as will be seen later, a consistent set is a weak equilibrium concept. It can be a nonsingleton set due to the relative ease in which deviations can be deterred. Along the dynamic dimension, farsightedness can easily lead to cycles of BA formation which are not observed in reality. The problem of cycles will be assumed away by assuming trade agreements are binding. As noted by Ornelas (2008), the assumption of binding agreements is a common implicit or explicit assumption in the literature that is also consistent with history.[7] A simple justification for this assumption is the existence of sufficiently large unmodelled costs of severance. As discussed by Baldwin (2008), tariffs are merely one element of the large and varied set of retaliation instruments, and the costs of severing BAs can far exceed that of lost exports.[8]

[6] The equilibrium concept used by Dutta et. al. (2005) is called an equilibrium process of network formation (EPNF). On one hand, an FDNE is more restrictive than an EPNF. While agreements are assumed to be binding in the FDNE, an EPNF allows each player in the randomly chosen "active pair" of a given period to sever any agreements it may have. Nevertheless, an EPNF does not allow players other than the active pair to sever existing agreements, which would be essential in the BA context if agreements were nonbinding. Moreover, owing to theoretical difficulties, Dutta et. al. (2005) assume that, other than the unilateral act of severing the agreement with the other member of the active pair, countries cannot condition upon unilateral acts.

[7] For example, it is an implicit assumption of Aghion et. al. (2007). There, the minimum transfer the leader country needs to offer the first follower country is the amount that leaves the follower country indifferent between the status quo and forming the BA. This is despite the fact that the follower may be subject to negative externalities when the leader subsequently forms a BA with the other follower and thus would want to sever its BA with the leader country.

[8] Retaliation instruments include the whole host of nontariff barriers to trade as well as development aid, military aid and political support in various international political issues. A breakdown in international cooperation could imply costs in the form of less progress in issues such as climate change, human rights, money laundering or illegal drugs. Further, as pointed out to me by Kamal Saggi, trade agreements often induce significant irreversible investment in physical capital on behalf of domestic firms, which would be severely devalued upon severance of the trade agreement.

Baldwin (2008, p.30) goes as far as to say "there is little to gain from explicitly modeling the threats since the most important forms of retaliation are entirely outside the model".

The third feature is that backward induction can be used to solve for the equilibrium. Practically, the equilibrium can be solved as follows. First, consider each hub–spoke network and solve for the simultaneous move equilibrium given global free trade will remain forever if reached. Second, move backward and consider each insider–outsider network. Now, solve for the simultaneous move equilibrium given how the network will evolve from each hub–spoke network. Third, move backward and consider the status quo network in which no agreements exist. Now, solve for the simultaneous move equilibrium given how the network will evolve from each insider–outsider network. The farsighted dynamic network equilibrium (FDNE) is then the sequence of networks on the equilibrium path. Essentially, there is no coalition who can feasibly alter the evolution of the network either on or off the equilibrium path and be better off.

After some basic terminology and notation in Section 2.2, Section 2.3 formally introduces the new dynamic network equilibrium concept of an FDNE.

2.2 Basic network terminology and notation

As argued by Jackson (2005, p.26–27), it is often easier in network and coalitional settings to completely dispense with strategies and focus directly on network stability. This is the approach adopted in this paper.[9] The terminology used closely follows Jackson (2008, especially pp. 376, 396–397) and, where relevant, Page et. al. (2005).

The finite set of nodes, or players, is $\mathcal{N} = \{1, 2, ..., n\}$. The game is an infinite horizon game with periods denoted by $t = 0, 1, 2, ...$. When nodes i and j are connected, there is a *link* between players i and j which is denoted by $\ell = ij$. In this paper, a link represents a trade agreement. A finite set of links is a *network*, $g = (\ell, \ell', \ell'', ...)$. Assume that links formed in previous periods cannot be severed and, at most, one link can form per period. Thus, the network in existence at the end of period t is $g_t = (\ell_1, \ell_2, ..., \ell_t)$ where ℓ_s is the link formed in period s. The complete network, g^c, is the network in which all possible links have formed. The set of possible networks is then $\mathcal{G}$, where $\mathcal{G} = \{g \mid g \subseteq g^c\}$. A *dynamic path* is an infinite sequence of networks $\{g, g', g'', ...\}$ with the set of possible dynamic paths being $\mathcal{P}$.

When link ℓ is added to network g, let the shorthand notation be that $g' = g + \ell$. Given g exists at the end of the previous period, the set of possible networks that can be formed in the current period is given by the correspondence $\mathcal{T}(g) = \{g' \mid g' = g + l, l \notin g \text{ or } \ell = \emptyset\}$ where $\mathcal{T} : \mathcal{G} \to \mathcal{G}$. Coalitions "act" by

[9] Goyal and Joshi (2006) and Furusawa and Konishi (2007) also adopt a network stability rather than strategy based approach. Moreover, even though Saggi and Yildiz (2010, p.29) use a strategy based equilibrium approach they nevertheless state "it proves convenient to refer directly to regime [network] changes rather than changes in announcements [strategies]". It is a trivial matter to reformulate and represent the FDNE in terms of strategies. However, the coalitional aspects of the framework make working with strategies more cumbersome.

making proposals and counter proposals from the set $\mathcal{T}(g)$. To this end, $T(g) = g'$ is a *g–proposal* where $T : \mathcal{G} \to \mathcal{T}(g)$. Given g is in existence at the end of the previous period, a g–proposal selects a network g' as the proposed network to be in place at the end of the current period. This selection comes from the set $\mathcal{T}(g)$. It is important to note that, by construction, g–proposals differ only in terms of the proposed link for the current period. Thus, the assumption that previously formed links cannot be severed is implicit in the definition of a g–proposal.[10]

Letting $\hat{T}(g) = \hat{g}'$ and $T(g) = g'$ be distinct g–proposals, Definition 2.1 now describes the feasibility of g–proposals.

Definition 2.1. The g–proposal $\hat{T}(g) = \hat{g}'$ is *obtainable* from $T(g) = g'$ by a coalition S if for $\ell = ij$:

i) $\ell \in g'$ and $\ell \notin \hat{g}'$ implies $i \in S$ or $j \in S$

ii) $\ell \in \hat{g}'$ and $\ell \notin g'$ implies $i \in S$ and $j \in S$

The interpretation of Definition 2.1 is straightforward. Part i) captures the aspect that link severance is a unilateral action; the link being removed from the g–proposal can be removed by either player in the link. Moreover, it can only be removed by such a player. It is important to stress that the link being severed here is not one that was previously formed; it is the proposed link to be formed in the current period. Severance and formation of this particular link is the essence of coalitions undertaking proposals and counter proposals. Part ii) captures the aspect that link formation is a bilateral action; the link being added to the g–proposal requires the consent of each player in the link.

When considering a coalitional deviation, the deviating coalition needs to know the sequence of networks that will follow whether the deviation occurs or not. To this end, let $M = \underset{g \in \mathcal{G}}{\cup} T(g)$ be a *proposal map* and the set of proposal maps be $\mathcal{M}$. Given any arbitrary network g that is in place at the end of the previous period, the proposal map specifies the proposed network that will be in place at the end of the current period. The proposal map is a crucial object. With it, the evolution of the network from the current period onwards can be traced out. Formally, given a proposal map M and a network g, $P(M, g) = \{g', g'', ...\}$ is the dynamic path which results from the current period onwards where $P : \mathcal{M} \times \mathcal{G} \to \mathcal{P}$.

Knowing the dynamic path from the current period onwards allows players, and coalitions, to compare continuation payoffs across various dynamic paths. Before defining the colaitional preference relation, individual payoffs are defined. Given g, let the one period payoff to player i be $\pi_i(g)$ where $\pi : \mathcal{G} \to \mathrm{I\!R}$. For the continuation payoff, it is useful to make the time notation explicit. Then, the continuation payoff player i receives from period t onwards given the dynamic path $P(M, g_{t-1}) = \{g_t, g_{t-1}, ...\}$ is $\Pi_i(\{g_t, g_{t-1}, ...\}) = \sum_{s=t}^{\infty} \beta^{s-t} \pi_i(g_s)$ where $\Pi : \mathcal{P} \to \mathrm{I\!R}$. For purposes of comparing dynamic paths,

[10]Formally, this can be seen because $T(g) \in \mathcal{T}(g)$ and, by construction, $g \subseteq g'$ for all $g' \in \mathcal{T}(g)$.

when the g–proposal $\hat{T}(g)$ replaces the g–proposal $T(g)$ in the proposal map M, the new proposal map is denoted by $\hat{M}$. Then, for a coalition $S \subseteq \mathcal{N}$, let $\succ_S$ be a relation on $\mathcal{P} \times \mathcal{P}$ where $P\left(\hat{M}, g\right) \succ_S P(M, g)$ iff $\Pi_i\left(P\left(\hat{M}, g\right)\right) > \Pi_i(P(M, g))$ for all $i \in S$. That is, a coalition prefers the dynamic path $P\left(\hat{M}, g\right)$ over the dynamic path $P(M, g)$ if and only if $P\left(\hat{M}, g\right)$ is associated with a higher continuation payoff for each player in the coalition.

2.3 Equilibrium concepts

2.3.1 The pure bilateralism game — "sequential link" equilibrium

In this subsection, the farsighted dynamic network equilibrium (FDNE) is developed for the case where previously formed links cannot be severed and, at most, one link forms per period. This corresponds to the pure bilateralism game. The extensions necessary to account for MFN agreements are dealt with in the following subsection. Definitions 2.2 and 2.3 place coalitional actions in the context of the coalitional simultaneous move game. Definitions 2.6, 2.7 and 2.8 embed the coalitional simultaneous move game in the dynamic framework.

Because players are farsighted, even in the simultaneous dimension, a coalition considers the sequence of subsequent deviations that its own coalitional deviation will induce. That is, it considers the sequence of counter proposals that its own counter proposal will induce in the given simultaneous move coalitional game. To this end, Definition 2.2 follows where the subscripts merely index elements of the sequence.

Definition 2.2. $T_0(g), T_1(g), ..., T_K(g)$ is a *simultaneous farsighted improving path* if for each $k = 1, ..., K$, the g–proposal $T_k(g)$ is obtainable from $T_{k-1}(g)$ by a coalition $S_k \subseteq \mathcal{N}$ such that $P(M_K, g) \succ_{S_k} P(M_{k-1}, g)$.

A simultaneous farsighted improving path is a K–sequence of obtainable coalitional deviations where each coalitional deviation is undertaken because, at the moment of the coalitional deviation, the final network in the sequence has a higher continuation payoff than the status quo for each member of the coalition. The equilibrium concept for the coalitional simultaneous move game is then as follows.

Definition 2.3. Given a proposal map M, a set of g–proposals is a *consistent set*, $\mathcal{C}(g)$, if for each g–proposal $T(g) \in \mathcal{C}(g)$ and for each $T_0(g)$ obtainable from $T(g)$ by a coalition S, either

i) $T_0(g) \in \mathcal{C}(g)$ and $P(M_0, g) \not\succ_S P(M, g)$ or

ii) there is a simultaneous farsighted improving path $T_0(g), ..., T_K(g)$ such that $T_K(g) \in \mathcal{C}(g)$ and $P(M_K(g)) \not\succ_S P(M, g)$

A set of g–proposals is consistent if, for each g–proposal in the consistent set, any obtainable deviation is deterred by another g–proposal in the consistent set. Deterrence can arise because the obtainable deviation is itself in the consistent set and is not preferred by the deviating coalition. It can also arise because the initial coalitional deviation induces a simultaneous farsighted improving path leading to another g–proposal in the consistent set that deters the initial coalitional deviation. It is important that deterrence arise because of another g–proposal in the consistent set because these are the only g–proposals which can be considered "stable". Given the possibility of consistent sets of different sizes, Chwe (1994) and Page et. al. (2005) argue the relevant consistent set is the largest consistent set. This is defined as follows and is natural given the union of consistent sets is itself consistent.

Definition 2.4. A consistent set of g–proposals, $\mathcal{C}(g)$, is the largest consistent set, denoted by $\mathcal{C}^*(g)$, if there is no other consistent set which is a superset of $\mathcal{C}(g)$.

The following definition is useful.

Definition 2.5. The g–proposal $T(g)$ is consistent if $T(g) \in \mathcal{C}^*(g)$.

The equilibrium concept of a consistent set is now embedded in a dynamic context.

Definition 2.6. $P(M, g_{t-1})$ is a *farsightedly dominant dynamic path* if the g–proposal $T(g_{s-1}) \in M$ is consistent for each $s \geq t$.

Definition 2.6 says that a dynamic path is farsightedly dominant from an arbitrary network if, given the proposal map M, there is no network along the dynamic path where a coalition could make a counter proposal so as to influence the evolution of the network and be better off. Requiring that a proposal map have this property from any network leads to the following definition.

Definition 2.7. The proposal map M is *farsightedly stable* if for each $g \in \mathcal{G}$, $P(M, g)$ is a farsightedly dominant dynamic path.

Thus, a farsightedly stable proposal map is one for which there is no network, on or off the equilibrium path, from which a coalition could make a counter proposal so as to alter the evolution of the network and be better off. The dynamic equilibrium is stated as follows.

Definition 2.8. The dynamic path P^* is a *farsighted dynamic network equilibrium* (FDNE) if there is a proposal map $M \in \mathcal{M}$ such that $P(M, \emptyset) = P^*$ and M is farsightedly stable.

Thus, an FDNE is merely a sequence of networks induced by a farsightedly stable proposal map. This subsection finishes with an important property of a farsightedly stable proposal map which will be frequently used. It assumes there are three players, which corresponds to the trade model to be introduced in the following section. Suppose that in each period, it is the preference of a pair of players to form a link between themselves or of a group of players, one from each link yet to form, to block any further links. Then the sequence of such links is the unique FDNE.

Lemma 2.1. *Assume there are three players and that M is farsightedly stable. Then,*

a) $\mathcal{C}^*(g) = T(g) = g'$ if $P(M,g) \succ_S P\left(\hat{M},g\right)$ for all $\hat{T}(g) \in \mathcal{T}(g)$ where $\hat{T}(g) \neq T(g)$ and either i) $S = \ell \neq \emptyset$ or ii) $\ell = \emptyset$ and $ij \subseteq S$ for any $ij \notin g$, and

b) $P(M,\emptyset)$ is the unique FDNE if for each $g \in \mathcal{G}$ on the equilibrium path, $\mathcal{C}^*(g)$ is a singleton.

Proof. See appendix.

2.3.2 The bilateralism and multilateralism games

The previous section laid out the pure bilateralism game; the links formed were BAs. For the bilateralism game, countries are also able to form MFN agreements where any preferential access granted to one country is granted to all. In what follows, attention will be restricted to the three country case. Thus, the possible MFN agreements are those between countries i and j, denoted $(ij)^{MFN}$, and that between countries i, j and k, denoted $(ijk)^{MFN}$. The presence of MFN agreements requires some adjustments to the notion of obtainability. The following definition extends Definition 2.1; parts i) and iii) of Definition 2.1 remain. Without loss of generality in the model of this paper, and merely for presentation purposes, it will be assumed that country k must abide by the MFN principle following $(ij)^{MFN}$ if it enters any agreements.

Definition 2.9. The g–proposal $\hat{T}(g) = \hat{g}'$ is *obtainable* from $T(g) = g'$ by a coalition S if conditions i)–ii) are satisfied as per Definition 2.1 where (ij) and $(ij)^{MFN}$ are treated as distinct links and

i) $g' = (ijk)^{MFN}$ implies $i \in S$ for some i and

ii) $\hat{g}' = (ijk)^{MFN}$ implies $i \in S$ and $j \in S$ for all $ij \notin g$ if $(ij)^{MFN} \neq g$ while $k \in S$ otherwise.

In addition to Definition 2.1, three adjustments are made. First, (ij) and $(ij)^{MFN}$ are treated symmetrically since they are distinct agreements. Second, any country can dissolve an MFN agreement made between the three countries. Third, implementing the three country MFN agreement requires the consent of all three countries unless there is an MFN agreement in place between countries i and j. In this case, the three country MFN agreement is reached if country k extends MFN access to both countries.

For the multilateralism game, Definition 2.9 applies directly except in that BAs are not permitted.

3 Underlying trade model

The underlying trade model is a three country oligopolistic intra industry trade model. Slight variations on this have been used by Krishna (1998), Ornelas (2005), Goyal and Joshi (2006), Mukunoki and Tachi

(2006) and Saggi and Yildiz (2011).[11] In this paper, the political economy approach is taken so government incentives for trade agreement formation depend only on firm profits. This extreme assumption sharpens the intuition behind the commitment problem and significantly increases the analytical tractability of the model. Moderate relaxation of this assumption would not affect the qualitative and intuitive results of the model because, as will be seen later, presence of the commitment problem merely requires that the largest country's payoff when an insider with the medium sized country exceed that under global free trade. Single period payoffs are realized at the end of each period through Cournot–Nash competition. Each country has a single representative firm with a common marginal cost which is normalized to zero. The inverse demand function in country j is $P^j = A^j - Q^j$ where the total quantity produced in country j is $Q^j = \sum_{k\in\mathcal{N}} q_k^j$. A^j captures the market size of country j and is the sense in which countries are asymmetric.[12]

A common nonprohibitive tariff, τ, is implemented by country j on the imports from country i if $ij \notin g$ where g represents the set of trade agreements in existence. Otherwise there is no tariff between i and j. Letting $\bar{\tau}$ be such that any tariff above $\bar{\tau}$ is prohibitive, $\tau_i^j(g) = \tau \leq \bar{\tau}$ if $ij \notin g$ while $\tau_i^j(g) = 0$ otherwise. Markets are assumed to be segmented which enables firms to price discriminate across countries. This and the political economy approach imply there is no tariff complementarity which in turn means the optimal tariff, subject to being non prohibitive, is $\bar{\tau}$.[13] The common nonprohibitive tariff and segmented markets ensure that each firm is active in each market.

A strand of the literature has viewed BAs being self enforcing via multilateral tariff cooperation as the central issue of BA negotiation (see Bagwell and Staiger (1997), Fruend (2000), Bond et. al. (2001) and Saggi (2006)). However, this emphasis on optimal self enforcing tariffs in an infinitely repeated game setting faces its own trade off. Baldwin (2008) has criticized this approach for two reasons. First, because tariffs are public information, the length of the cheating phase is arbitrarily short meaning the one shot benefit from deviation is arbitrarily small.[14] Second, tariffs are merely one element of the set of retaliation instruments, so severance costs go far beyond that incurred through tariff retaliation. As such, while the assumption of a common tariff abstracts from a rigorous determination of optimal tariffs, it avoids these conceptual problems and provides significant analytical tractability in a complex dynamic setting.

[11] Here, the reduced form specification is presented. See Ornelas (2005, pp.475–476) for a detailed description of the structural assumptions justifying this reduced form.

[12] Letting c^j be the marginal cost of the firm from country j, $A^j - c^j$ is the important term governing asymmetry in a model such as this. As such, some papers hold A^j constant across countries while varying c^j. Others do the converse, and some directly focus on varying $A^j - c^j$ across countries.

[13] It is easy to verify that $\bar{\tau} = \frac{A^s}{3}$.

[14] Baldwin (2008) describes the surprise announcement by U.S. President Nixon in 1971 of a uniform 10% rise in U.S. tariffs; the immediate foreign reaction precluded implementation of the tariff increases.

Given g, the firm from country i maximizes profits by solving the following problem in each country j:

$$\max_{q_i^j} q_i^j \left[A^j - Q^j - \tau_i^j(g)\right].$$

Equilibrium quantity and profit for this firm are $q_i^{j*}(g) = \frac{1}{4}\left[A^j + \sum_{k \notin N^j} \tau_k^j - 4\tau_i^j(g)\right]$ and $\pi_i^j(g) = \left(q_i^{j*}(g)\right)^2$ where N^j is the set of countries who have a BA with country j (including country j itself). The payoff for the country j firm is then $\pi_i(g) = \sum_{j \in \mathcal{N}} \pi_i^j(g)$.[15] Country i's continuation payoff from period t onwards is then $\Pi_i(g_t) = \sum_{s=t}^{\infty} \beta^{s-t} \pi_i(g_s)$ where $\beta \in (0,1)$ is the common discount factor. If the network remains unchanged forever from period T onwards, $\pi_i(g_t) = \pi_i(g_T)$ for all $t \geq T$.

The key static incentives in the model revolve around the net benefit of forming a BA and the negative externalities that third party countries suffer. When country i forms a BA with country j, the discrimination country i faced in country j disappears, and its profits rise by $\frac{3\tau}{16}\left(2\hat{q}_i^{j*}(g) + 3\tau\right)$ where $\hat{q}_i^{j*}(g) \equiv \left(q_i^{j*}(g)\right)^{\frac{1}{2}} > 0$. This is the benefit of preferential access for country i. It is also increasing in the market size of country j since $\frac{\partial q_i^{j*}(g)}{\partial A^j} > 0$; preferential access is more valuable in a larger market.

However, the preferential access given by country i to country j lowers the profits of both country i and k by $\frac{\tau}{16}\left(2\hat{q}_c^{j*}(g) - \tau\right)$ for $c = i, k$. For country i, this is the cost of preferential access incurred when entering a BA with country j. For country k, it is the third party negative externality it faces in country i. This effect is also increasing in the market size of country i since $\frac{\partial q_c^{i*}(g)}{\partial A^i} > 0$. As before, the preferential access gained by country j is more valuable in a larger market which makes it more costly for countries i and k. These static incentives imply a country prefers to form a BA a larger partner. This is summarized in Lemma 3.1.

Lemma 3.1. *The one period change in profits for country i when forming a BA with country j is increasing in $A^j - A^i$.*

Proof. Since $\frac{\partial q_i^{j*}(g)}{\partial A^i} = \frac{\partial q_i^{i*}(g)}{\partial A^j} = 0$, this follows from the benefit of preferential access being increasing in A^j on account of $\frac{\partial q_i^{j*}(g)}{\partial A^j} > 0$ and the cost of preferential access being decreasing in A^i on account of $\frac{\partial q_i^{i*}(g)}{\partial A^i} < 0$.■

While a country prefers to form a BA with a larger partner, the net benefit for country i when entering a BA with country j is $\Delta\pi_i^j(g) = \frac{3\tau}{16}\left(2\hat{q}_i^{j*}(g) + 3\tau\right) - \frac{\tau}{16}\left(2\hat{q}_i^{j*}(g) - \tau\right)$. This reduces to $\Delta\pi_i^j(g) \propto 2\hat{q}_i^{j*}(g) + \tau + \left(A^j - A^i\right) + \tau\left(p^i(g) - p^j(g)\right)$ where $p^c(g)$ for $c = i, j$ is the number of BAs that country c has (including with itself) given network g. In addition to Lemma 3.1, notice that $\Delta\pi_i^j(g)$ is increasing in $p^i(g) - p^j(g)$; a greater number of existing partners reduces the cost of additional preferential access because the degree of

[15] Network dependent one period payoffs are derived in Appendix A.

protection is lower. In addition, regardless of the trade network, this leads to the following implications.

Lemma 3.2. *The one period change in profits is positive when two symmetric countries form a BA. Regardless of asymmetry, it is positive for the smaller country.*

Proof. Under symmetry, the lower bound of $\Delta\pi_i^j(g)$ occurs for $p^j(g) = 2$ and $p^i(g) = 1$. Hence, the lower bound is $\Delta\pi_i^j(g) \propto 2\hat{q}_i^{j*}(g) > 0$. Since $\Delta\pi_i^j(g) > 0$ when $A^i = A^j$ and $\Delta\pi_i^j(g)$ is increasing in $A^j - A^i$, then $A^j > A^i$ implies $\Delta\pi_i^j(g) > 0$.■

Thus, from a one shot perspective, any country finds it profitable to form any BA under symmetry while only the larger country may find it unprofitable under asymmetry.

4 Equilibrium of the pure bilateralism game

4.1 Symmetric market size

The case of symmetry will be developed initially to help provide general intuition before introducing asymmetry. At the heart of this intuition is the commitment problem faced by insiders, and underlying this commitment problem when countries are symmetric is Lemma 3.2. While the extent to which a country's profits rise upon BA formation depends on how many existing BAs each of the partners has and the market size of each partner, profits rise nonetheless under symmetry.

Notationally, $\pi_i^{p,j}$ will denote country i's profit given occupation of network position p with country j. Figure 1 shows each network position and its notation. Given Lemma 3.2, there is an incentive for an insider to form a BA with the outsider and become the hub since $\pi_i^H > \pi_i^{I,j}$. However, when either insider becomes the hub, the two spokes will subsequently form their own BA as $\pi_i^{FT} > \pi_i^{K,j}$. Moreover, $\pi_i^{I,j} > \pi_i^{FT}$ implies lost insider rents and negative third party externalities to do not compensate for the removal of discrimination in market k.[16] Thus, an insider faces a trade off. While being the hub entails higher rents than being an insider due to the additional preferential access, free trade entails lost insider rents. An insider will prefer to become the hub rather than remain an insider when it is sufficiently impatient as it then places sufficient value on the additional rents earned as the hub. The formal statement of this is $\pi_i^H + \frac{\beta}{1-\beta}\pi_i^{FT} > \frac{1}{1-\beta}\pi_i^{I,j}$ or $(1-\beta)\left(\pi_i^H - \pi_i^{I,j}\right) + \beta\left(\pi_i^{FT} - \pi_i^{I,j}\right) > 0$. Let $\sigma \equiv (\alpha_{ls}, \alpha_{ms}, \tau)$ where $\alpha_{ls} = A^l/A^s$ and $\alpha_{ms} = A^m/A^s$ are the relative market sizes of l and m with respect to s and, abusing notation, τ is now the original tariff

[16] In general, as shown in Appendix A, $\pi_i^{I,j} - \pi_i^{FT} \propto A^i + A^j - 2A^k + 3t$. This is positive under symmetry for any i and always positive when $k = s$ meaning that l and m are insiders.

normalized by A^s. The previous inequality then reduces to the free trade–insider condition or FT–I condition

$$\beta < \bar{\beta}_{i,j}^{FT-I}(\sigma) \equiv \frac{\pi_i^H - \pi_i^{I,j}}{\pi_i^H - \pi_i^{FT}}. \tag{1}$$

When the FT–I condition is satisfied, both insiders prefer to remain insiders over becoming the hub on a path to free trade.[17] Nevertheless, and here is the crux of the commitment problem, *if* one of them believes the other will become the hub then it also has an incentive to become the hub as a means to avoid becoming a spoke since $\pi_i^H > \pi_i^{K,j}$ and free trade would be reached anyway. The key question is then whether the insiders can commit to not become the hub.

Under symmetry, the answer is simple and is depicted in Figure 2. When the FT–I condition holds, both insiders recognize that becoming the hub would not only hurt the other insider but also itself. Thus, each can credibly commit to not doing so. Initially, it may appear that the commitment problem is not binding once the FT–I condition is satisfied since both insiders prefer to become the hub on the path to free trade rather than remain an insider. However, under symmetry, the outsider is indifferent as to which insider it forms a BA with. As will be seen soon, this leads to multiple equilibria; each insider becoming the hub is an equilibrium. If each insider views becoming the hub or the spoke as equally likely events, then it would prefer to remain an insider if $\frac{1}{1-\beta}\pi_i^{I,j} > \frac{1}{2}\left(\pi_i^H + \frac{\beta}{1-\beta}\pi_i^{FT}\right) + \frac{1}{2}\left(\pi_i^{K,k} + \frac{\beta}{1-\beta}\pi_i^{FT}\right)$. Simple algebra reveals this holds given $\beta < \bar{\beta}^{FT-I}(\sigma)$. Thus, ex ante, the insiders still face a commitment problem as both would be better off if they could commit to remain insiders. However, since the FT–I condition holds, they cannot commit to remain insiders; each prefers the path to free trade on which it is the hub over remaining an insider.

Given the intuition underlying the commitment problem, the following proposition is quite intuitive and is depicted in Figure 3. Henceforth, to simplify notational clutter although it is an abuse of notation, (ij, ik, jk) will represent the sequence of networks obtained by i and j forming the BA in period one followed by i and k in period two and j and k in period three. Finally, given an FDNE, BAs are said to be "building blocs" if global free trade is attained and "stumbling blocs" otherwise.

Proposition 4.1. *Under symmetry, BAs are stumbling blocs when insiders can commit to not become the hub; any insider–outsider network is an FDNE. Otherwise, BAs are building blocs; any path to free trade is an FDNE. There are no other FDNE.*

When insiders can overcome the commitment problem and commit to remain insiders, i.e. the FT–I condition holds, then an insider–outsider network will form in equilibrium. In Figure 3, the set of such

[17] Under symmetry, the FT–I condition reduces to $\bar{\beta}^{FT-I}(\tau) = \frac{2(1-\tau)}{2+\tau}$. Given the nonprohibitive traiff condition and $\tau > 0$, $\bar{\beta}^{FT-I}(\tau) \in (0, 1)$.

networks is denoted Θ^{I-O}. The multiplicity of equilibria is due to symmetry. Symmetry implies each country is indifferent to whom it becomes an insider with meaning that any insider–outsider network could emerge in equilibrium. The presence of asymmetry will enable prediction of which insider–outsider network will emerge. When the FT–I condition does not hold, then an insider prefers to be a hub on the path to free trade rather than remain an insider; the additional preferential access and rents as the hub more than compensate for the elimination of rents in free trade. Here, any path to free trade could emerge in equilibrium. In Figure 3, the set of such paths is denoted Θ^{FT} and, again, symmetry drives the multiplicity of equilibria. The status quo of no BAs cannot emerge in equilibrium despite not knowing which path will lead to free trade because, ex ante, the expected profit across all paths exceeds the status quo profit.

4.2 Asymmetric market size — "similar" countries

This section moves away from the symmetric case by introducing a moderate amount of market size asymmetry. Introducing asymmetry complicates matters because it means that different countries face different incentives for BA formation depending on the specific BA partner. In particular, there is a distinct commitment problem between each pair of insiders. Moreover, for a given pair of insiders, different incentives shape each insider's commitment problem. After developing an understanding of the commitment problem in the presence of asymmetry, Section 4.3 will introduce a greater degree of asymmetry with dramatic effects for the equilibrium.

What is meant by a moderate degree of asymmetry will now be formalized. Two countries are said to be *similar*, denoted $i \approx j$, if the following is true from a myopic perspective and from any network: any two countries prefer to form a BA between themselves if it may actually form but has not. The qualifier "may actually form" is merely used to exclude the BA between s and l when s and m are insiders. In this case, from a myopic perspective, the BA between s and l would never form. By Lemma 3.1, if l benefits from forming a BA with s, then it prefers to form the BA with m and m is willing to do so. The conditions for countries to be similar are formalized in Lemma 4.1.

Lemma 4.1. *l and s are similar, denoted $l \approx s$, if $\pi_l^{FT} > \pi_l^{K,s}$. m and s are similar if $\pi_m^{FT} > \pi_m^{K,s}$. l and m are similar if $\pi_l^{K,s} > \pi_l^{O}$. These conditions, respectively, reduce to $\alpha_{ls} < 3 - 5\tau$, $\alpha_{ms} < 3 - 5\tau$ and $\alpha_{ls} < 3\alpha_{ms} - 6\tau$.*

The importance of countries being similar for the commitment problem is that it implies any pair of spokes will form a BA, which is a key ingredient underlying the commitment problem.[18] Given this, the key ingredient determining whether an insider prefers to become the hub is whether, relative to being an insider,

[18]Note that, under this definition, there is a small range of the parameter space for which $l \approx s$ but $l \approx m$ does not hold. This part of the parameter space is uninteresting and, as such, will be ignored.

the additional preferential access from being a hub outweighs the loss of preferential access upon free trade. Using the FT–I condition, an insider prefers to become the hub when $(1-\beta)\left(\pi_i^H - \pi_i^{I,j}\right) + \beta\left(\pi_i^{FT} - \pi_i^{I,j}\right) > 0$. Since both insiders have the same degree of preferential access in each of the three markets under free trade and as insiders, then $\left(\pi_i^{FT} - \pi_i^{I,j}\right) = \left(\pi_j^{FT} - \pi_j^{I,i}\right)$. While, in addition, the additional preferential access earned as a hub is identical for either insider, the cost incurred by an insider in becoming the hub is different. The cost is larger for the larger insider. As such, the smaller insider has a greater incentive to become the hub. Letting $\bar{\beta}_{i,j}^{FT-I}(\sigma)$ be i's FT–I condition as an insider with j and letting $A^i > A^j$, this implies $\bar{\beta}_{i,j}^{FT-I}(\sigma) < \bar{\beta}_{j,i}^{FT-I}(\sigma)$ as illustrated in Figure 4.

As in the case of symmetry, the insiders can commit to remain insiders when neither insider's FT–I condition is satisfied; the negative externalities suffered and rents lost are larger than the benefits of being the hub. Under asymmetry, this occurs for $\beta > \bar{\beta}_{j,i}^{FT-I}(\sigma) > \bar{\beta}_{i,j}^{FT-I}(\sigma)$. In this case, conditional on i and j being insiders, they will remain so. Unlike the case of symmetry, the commitment problem is not binding when the FT–I conditions of both insiders hold. Because the larger insider, i, offers greater preferential access to the outsider, k, then the larger insider is the hub on the path to free trade. Another difference to the symmetric case is an intermediate range of β for which either insider may become the hub or they remain as insiders; the largest consistent set in period two is $\mathcal{C}(ij) = \{(ij), (ij, ik), (ij, jk)\}$. Whether the commitment problem is still binding, in which case it cannot be overcome since $\mathcal{C}(ij) \neq (ij)$, depends on whether the insiders prefer to remain insiders rather than face a uniform distribution over $\mathcal{C}(ij)$. As insiders, m and l have the greatest degree of preferential access to protect, and their commitment problem remains binding. For s and m it is not binding, while for s and l it depends on the degree of asymmetry.[19]

The essence of the multiplicity of equilibria is the farsightedness of countries. Figure 5 illustrates this issue, remembering that the spokes will form the final BA. $\mathcal{C}(ij)$ contains each possible network in period 2 because any coalitional deviation from any network which makes the coalition immediately better off is deterred by a subsequent deviation to another "stable" network. Stable in the sense that it is also in the consistent set. To see this, start with the network in which the insiders commit to remain insiders. While $\beta < \bar{\beta}_{j,i}^{FT-I}(\sigma)$ implies both j and k would benefit from j becoming the hub, i would be the spoke. Since i provides greater preferential access to k, i and k would then make a counter proposal in which i is the hub. This deters k from being involved in the initial deviation. Nevertheless, the network involving j as the hub is also in the consistent set. While k would benefit from the aforementioned counter proposal, i and j would then benefit from remaining insiders and would make a counter proposal to do so. This deters k from making the counter proposal with i. Finally, the network in which i is the hub is also stable despite i and j both preferring to remain insiders. i is deterred from making such a counter proposal deviation because j

[19]Simple algebra reveals the condition determining whether the commitment problem is binding is $A^i + A^j - 2A^k > 0$.

and k would then benefit from making j the hub.

While asymmetry complicates matters since the commitment problem between each pair of insiders is characterized by different incentives, fortunately the commitment problem faced by m and l as insiders is the crucial one for the analysis. This is because if m and l can commit to remain insiders, they prefer this over becoming the hub on the path to free trade and thus over any path on which they are an insider with s. This merely follows from the FT–I condition and the fact that being an insider with a larger partner offers greater preferential access. This primacy of the commitment problem between m and l underlies the following proposition which is illustrated in Figure 6.

Proposition 4.2. *When countries are similar and m and l can commit to not become the hub, they remain insiders and BAs are stumbling blocs. Otherwise, there may be multiple FDNE. If so, there is at most one FDNE in which BAs are stumbling blocs. The only unique building bloc FDNE occurs for $\beta < \bar{\beta}^{m}(\sigma) \equiv \frac{\pi_m^{I,l} - \pi_m^{I,s}}{\pi_m^{H} - \pi_m^{K,s}}$ where the path to free trade is (ml, sl, sm).*

Figure 6 shows that when m and l can overcome the commitment problem, they remain insiders, which implies BAs are stumbling blocs to free trade. For $\beta \in \left(\bar{\beta}_{l,m}^{FT-I}(\sigma), \bar{\beta}_{m,l}^{FT-I}(\sigma)\right)$, there is a multiplicity of equilibria stemming from an interaction between the inability of m and l to overcome the commitment problem and the farsightedness of countries as described above and illustrated in Figure 5. Thus, each of (ml), (ml, sm, sl) and (ml, sl, sm) is an FDNE. These three paths constitute Θ_1^{FT} in Figure 6.

In contrast, the multiplicity of equilibria for $\beta \in \left(\bar{\beta}^{m}(\sigma), \bar{\beta}_{l,m}^{FT-I}(\sigma)\right)$ arises from a different source. Here, the commitment problem for m and l is not binding because l is willing to become the hub despite the negative externalities that will follow, and s, as a spoke, prefers the larger insider to be the hub. Here, the source of the multiplicity is that each country prefers a different pair of insiders on the path to free trade.[20] Knowing it will be the hub, l prefers to be an insider with m. However, in order to be the hub, m prefers to be an insider with s so it can be the hub. Knowing it will be a spoke, s prefers to be an insider with l. Similar logic to the previous paragraph implies each such outcome is an FDNE. Any deviation which appears beneficial is deterred because of fears about subsequent deviations by other countries. These three paths to free trade together with the status quo of no BAs constitute Θ_2^{FT} in Figure 6.

However, $\bar{\beta}^{m}(\sigma)$ proves to be a critical value below which uniqueness of the equilibrium is restored. This is because m faces a trade off when comparing the path on which it is an insider with s to the path on which it is an insider with l. While the preferential access embodied in being an insider with l exceeds that of being an insider with s, m will only be the hub if it is an insider with s. When the discount factor is sufficiently small, $\beta < \bar{\beta}^{m}(\sigma)$, the benefit of being an insider with l dominates, and both m and l mutually prefer to be

[20] Remember that $\beta < \bar{\beta}_{l,m}^{FT-I}(\sigma)$ implies $\beta < \bar{\beta}_{i,j}^{FT-I}(\sigma)$ for any i, j so that free trade will arise conditional upon any pair of insiders.

insiders on the path to free trade. With a general characterization of the equilibria for similar countries, it is now possible to see how this characterization depends on the degree of asymmetry.

Proposition 4.3. *When countries are similar, greater asymmetry increases the extent to which m and l can overcome the commitment problem. This increases the extent to which BAs are stumbling blocs.*

The essence of Proposition 4.5 is illustrated, without loss of generality, in Figure 7 for the case of $\tau = \frac{1}{4}$. Notationally, $\alpha_{ms} = A^m/A^s$ and $\alpha_{ls} = A^l/A^s$ are the relative market sizes of m and l with respect to s. The curves labelled $\bar{\alpha}_{ms}$ are contour curves showing, for a given α_{ms}, the pairs of $\left(\alpha_{ls}, \bar{\beta}_{m,l}^{FT-I}(\sigma)\right)$ such that m's FT–I condition with l is satisfied for $\beta < \bar{\beta}_{m,l}^{FT-I}(\sigma)$.

Because the $\bar{\alpha}_{ms}$ contour curves are downward sloping, Figure 7 illustrates that greater asymmetry from a higher α_{ls} or α_{ms} increases the range of β for which m's FT–I condition with l fails. Thus, greater asymmetry increases the extent to which m and l can overcome their commitment problem and, in turn, the extent to which BAs are stumbling blocs. With respect to m (similar logic applies to l), a higher α_{ls} increases the incentive to remain an insider. It does so by increasing the rent associated with being an insider, which would be lost on the path to free trade. A higher α_{ms} also increases the incentive for m to remain an insider, but does so because it increases the cost of preferential access m would incur by becoming the hub.

This section has showed that the extent to which BAs are stumbling blocs is related to the ability of m and l to overcome their commitment problem and greater asymmetry increases this ability. However, the degree of asymmetry considered so far was moderate enough that l would still find it beneficial to form a BA with s if they were both spokes. This was central to the commitment problem. The next section relaxes this assumption with dramatic effects on the equilibrium because the commitment problem between m and l breaks down.

4.3 Asymmetric market size — "nonsimilar" countries

The critical part of Lemma 4.1, which defines countries as being similar, is that, as a spoke, l would form the BA with s . Now suppose this is not so. Then, as an insider with l, m faces no negative externalities by becoming the hub because it would remain the hub forever. Thus, m cannot credibly commit to not become the hub.

The first step in understanding the effect of the commitment problem breakdown is understanding the new trade off faced by l as an insider with m. Since m cannot commit to remain an insider, l can either let m become the hub knowing that it will remain a spoke or it can become the hub knowing that free trade will follow. This trade off is captured by the free trade–spoke condition, or the FT–K condition. To focus on this trade off, assume that $\pi_l^H > \pi_l^{I,m}$, or $\alpha_{ls} \leq 3 - 2\tau$, which ensures l benefits from becoming the hub. l is then *big* relative to s if $\pi_l^{K,s} \geq \pi_l^{FT}$ and $\pi_l^H > \pi_l^{I,m}$.[21] l will prefer to become the hub when $\pi_l^H + \frac{\beta}{1-\beta}\pi_l^{FT} > \frac{1}{1-\beta}\pi_l^{K,s}$ or $(1-\beta)\left(\pi_l^H - \pi_l^{K,s}\right) + \beta\left(\pi_l^{FT} - \pi_l^{K,s}\right) > 0$. This reduces to

$$\beta < \bar{\beta}^{FT-K}(\sigma) \equiv \frac{\pi_l^H - \pi_l^{K,s}}{\pi_l^H - \pi_l^{FT}}. \tag{2}$$

Because l prefers to remain a spoke with s over free trade, it must be sufficiently impatient for the benefit of becoming the hub to dominate.

The second step in understanding the effect of the commitment problem breakdown starts with the fact that the commitment problem was easier to overcome for m and l than l and s when countries were similar, i.e. $\bar{\beta}_{l,m}^{FT-I}(\sigma) < \bar{\beta}_{l,s}^{FT-I}(\sigma)$. As such, l preferred to be an insider with m rather than s. Even if the commitment problem between l and m could not be overcome or was not binding, l still preferred m as its fellow insider due to the higher rents on the path to free trade. However, breakdown of its commitment problem with m means l may prefer to be an insider with s if they can overcome their commitment problem as insiders; i.e. $\beta < \bar{\beta}_{s,l}^{FT-I}(\sigma)$. While l would forego the higher rents associated with having m as its fellow insider, it would keep its rents as an insider with s rather than losing its insider rents when m becomes the hub. The following proposition provides an initial characterization of when BAs are building blocs.

Proposition 4.4. *Suppose s and l cannot overcome their commitment problem and l prefers to become the hub after being an insider with m, i.e. $\beta < \bar{\beta}^{min}(\sigma) \equiv \min\left\{\bar{\beta}^{FT-K}(\sigma), \bar{\beta}_{s,l}^{FT-I}(\sigma)\right\}$. If there are multiple FDNE, there are FDNE in which BAs are building blocs. Moreover, BAs are building blocs in the only unique FDNE. This FDNE is (ml, sl, sm) which, given critical values $\bar{\beta}^s(\sigma)$ and $\bar{\beta}^m(\sigma)$, arises for $\beta < \bar{\beta}^m(\sigma)$ and $\beta > \bar{\beta}^s(\sigma)$.*

Since m faces no negative externalities by becoming the hub when it is an insider, l's decision about whether to become the hub is driven by the FT–K condition. Once this is satisfied and the commitment problem between s and l cannot be overcome, there is an equilibrium in which free trade is reached. The potential multiplicity of equilibria between the critical values $\bar{\beta}^m(\sigma)$ and $\bar{\beta}^s(\sigma)$ arises because each country prefers a different dynamic path for the global trade network. No country can build a coalition to ensure its preferred dynamic path. In the standard case, similar to Proposition 4.2 for $\beta \in \left(\bar{\beta}^m(\sigma), \bar{\beta}_{l,m}^{FT-I}(\sigma)\right)$, l wants to be an insider with m and then the hub while m wants to be an insider with s so it can then be the hub. In contrast, s wants to be an insider with l because, knowing it cannot be the hub, it prefers the larger insider rents with l.[22] Each of these dynamic paths is an equilibrium because any coalitional deviation that

[21] Hence, this section assumes that $\alpha_{ls} \in [3-5\tau, 3-2\tau)$ since $\pi_l^{K,s} \geq \pi_l^{FT}$ reduces to $\alpha_{ms} \geq 3-5\tau$. It also assumes that $l \approx m$ which, for $\alpha_{ls} \leq 3 - c\tau$ where $c \in (2,5]$, implies $\alpha_{ms} \geq 1 + \frac{1}{3}(6-c)\tau$.

[22] The proof of Proposition 4.4 details all possible cases of multiplicity.

seems beneficial is deterred by the fear of subsequent coalitional deviations that harm the initial deviating coalition. Indeed, even the status quo of no BAs can be sustained as an equilibrium because of these fears.

However, for $\beta < \bar{\beta}^{m}(\sigma)$ and $\beta > \bar{\beta}^{s}(\sigma)$, a coalition forms to ensure (ml, sl, sm) is the unique FDNE. Given l's preference to become the hub, m and l form the coalition when m is sufficiently impatient that the attractiveness of insider rents with l rather than s offsets the foregone rents it would earn as a hub if it became an insider with s. In contrast, s and l may form the coalition. Here, it is important that s will remain a spoke forever after being an insider with m because l will not form a BA with s. Then s and l form the coalition when s is sufficiently patient that the benefit of forming BAs with both m and l outweighs the foregone rents as an insider with m. In this situation, l and s can credibly threaten to become insiders themselves. This induces m to agree to be an insider with l even though it would prefer to be an insider with s.

The breakdown in the commitment problem between m and l dramatically alters the extent to which BAs are building blocs to free trade. The switch from the FT–I condition to the FT–K condition as the characterization of the dynamic incentives facing m and l as insiders is the primary cause. When the FT–I condition was satisfied, l and m refrained from becoming the hub because they knew the other would also do so. However, under the FT–K condition, l knows that m will not refrain from becoming the hub so l also exercises less restraint. The following proposition details the increased extent to which BAs are building blocs.

Proposition 4.5. *Breakdown of the commitment problem between m and l causes a dramatic once off jump in the extent to which BAs are building blocs. Formally, fixing $\bar{\alpha}_{ms} < 3 - 5\tau$ and $\bar{\alpha}_{ls} = 3 - 5\tau$, $\bar{\beta}^{min}(\hat{\sigma}) > \bar{\beta}^{FT-I}_{m,l}(\bar{\alpha}_{ms}, \bar{\alpha}_{ls}, \hat{\tau})$ for any $\hat{\sigma}$. Conditional on this breakdown, greater asymmetry between l and s increases the extent to which BAs are stumbling blocs, while greater asymmetry between m and s increases the extent to which BAs are building blocs.*

The basic results can be seen, without loss of generality, in Figure 8 where $\tau = \frac{1}{4}$. When l becomes big relative to s, i.e. $\alpha_{ls} > 3 - 5\tau = 1.75$, the $\bar{\alpha}_{ms}$ contour curves depict the critical value $\bar{\beta}^{min}(\sigma)$. With l big relative to s, the $\bar{\alpha}_{ms} = \underline{\alpha}_{ms}$ curve represents the binding constraint on α_{ms} to ensure l and m remain similar.

Conditional on the commitment problem breakdown between m and l, i.e. $\alpha_{ls} > 1.75$, Figure 8 clearly illustrates the effect of asymmetry. A higher α_{ls} is a movement down an $\bar{\alpha}_{ms}$ contour curve because it increases both the cost of preferential access incurred by l when becoming the hub and the value to s of preserving insider rents with l. Respectively, these reduce the incentive for l and s to become the hub which reduces both $\bar{\beta}^{FT-K}(\sigma)$ and $\bar{\beta}^{FT-I}_{s,l}(\sigma)$. Conversely, a higher α_{ms} increases the rents associated with becoming the hub by making preferential access with m more valuable. The greater incentive for s and l to become the hub raises $\bar{\beta}^{FT-K}(\sigma)$ and $\bar{\beta}_{s,l}^{FT-I}(\sigma)$. Thus, a higher α_{ms} shifts the $\bar{\alpha}_{ms}$ contour curve upwards.

That a higher α_{ms} increases the extent to which BAs are building blocs is the opposite of when l and s were similar. When l and s were similar, free trade was reached when l was willing to *give up* sole preferential access with m as an insider by becoming the hub and accepting the subsequent negative externalities. However, now free trade is reached once s and l are sufficiently willing to *have* sole preferential access with m by becoming the hub. This dramatic change arises because the dynamic incentive structure is now characterized by the FT–K condition and s's FT–I condition with l rather than m's FT–I condition with l.

The extent to which the commitment problem breakdown increases the role of BAs as building blocs can be seen by the fact that that along the $\bar{\alpha}_{ms} = \underline{\alpha}_{ms}$ contour curve , $\bar{\beta}^{min}(\sigma)$ is greater than $\bar{\beta}_{m,l}^{FT-I}(\sigma)$ on any $\bar{\alpha}_{ms}$ contour curve for $\alpha_{ls} = 3 - 5\tau$. Thus, the small change in asymmetry leading to l being big relative to s significantly increases the extent to which BAs are building blocs. Again, the intuition revolves the change in the dynamic incentive structure. Even though the relevant inequality of interest is really $\bar{\beta}^{min}(\sigma) = \min\left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}_{s,l}^{FT-I}(\sigma)\right) > \bar{\beta}_{m,l}^{FT-I}(\sigma)$ rather than $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}_{m,l}^{FT-I}(\sigma)$, the fact that the cost of preferential access is lower for s than m implies $\bar{\beta}_{s,l}^{FT-I}(\sigma) > \bar{\beta}_{m,l}^{FT-I}(\sigma)$. Therefore, the critical inequality is $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}_{m,l}^{FT-I}(\sigma)$ which is directly related to breakdown of the commitment problem between m and l. When l and s were similar, l and m could credibly commit to remain insiders because each knew of the subsequent negative externalities it would face when the two spokes formed their own BA. However, because m no longer faces this fear when l is big relative to s, m cannot credibly commit to not become the hub. This commitment problem breakdown forces l to become the hub for an expanded range of β and explains why $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}_{m,l}^{FT-I}(\sigma)$ at the critical value of $\alpha_{ls} = 3 - 5\tau$ and thus explains the associated jump in the extent to which BAs are building blocs.

Propositions 4.4 and 4.5 have focused on the role of BAs as building blocs, but have not characterized the equilibrium when $\beta > \bar{\beta}^{min}(\sigma)$. Proposition 4.6 and Figure 9 now do so. For the purposes of Proposition 4.6, let $\bar{\beta}^{l}(\sigma)$ be such that l prefers the dynamic path resulting from (sl) rather than (ml) for $\beta > \bar{\beta}^{l}(\sigma)$.

Proposition 4.6. *Suppose $\beta \geq \bar{\beta}^{min}(\sigma)$ and l prefers remaining an insider with s over the status quo of no BAs. For $\bar{\beta}^{min}(\sigma) = \bar{\beta}^{FT-K}(\sigma)$, the FDNE is (sl) for $\beta \in \left(\bar{\beta}^{l}(\sigma), 1\right)$ and (ml, sm) for $\beta \in \left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}^{l}(\sigma)\right)$. For $\bar{\beta}^{min}(\sigma) = \bar{\beta}_{s,l}^{FT-I}(\sigma)$, the FDNE is (sl) for $\beta \in \left(\bar{\beta}^{l}(\sigma), 1\right)$, (ml, sm) for $\beta \in \left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}^{l}(\sigma)\right)$ and (ml, sl, sm) for $\beta \in \left(\bar{\beta}_{s,l}^{FT-I}(\sigma), \min\left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}^{l}(\sigma)\right)\right)$.*

Figure 9 illustrates 4.6. Given the BA between s and l will not arise when they are spokes, (sm, ml) will result conditional on s and m being insiders. However, the FT–I and FT–K conditions still govern network formation conditional on the other insider–outsider networks.

Only for very large degrees of asymmetry is it true that $\bar{\beta}^{FT-I}_{s,l}(\sigma) > \bar{\beta}^{FT-K}(\sigma)$. The top pane of Figure 9 depicts this case. Since it can be shown that $\bar{\beta}^{l}(\sigma) > \bar{\beta}^{FT-I}_{s,l}(\sigma)$, $\bar{\beta}^{l}(\sigma)$ splits the parameter space as depicted since s prefers (sl) and m prefers (ml, sm) to any other feasible networks. Then, free trade is attained once l prefers (ml, sl, sm) over (ml, sm), which happens for $\beta < \bar{\beta}^{min}(\sigma) = \bar{\beta}^{FT-K}(\sigma)$. If α_{ms} is sufficiently large, the net benefit l receives from a BA with s is small enough that l would prefer to be a spoke under (ml, sm) rather than an insider with s regardless of β. Then, (sl) does not exist in equilibrium.

The bottom panes of Figure 9 depict the standard case of $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}^{FT-I}_{s,l}(\sigma)$. The rather simplistic equilibria structure for the case of $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}^{FT-I}_{s,l}(\sigma)$ becomes more complicated because $\bar{\beta}^{l}(\sigma) \gtrless \bar{\beta}^{FT-K}(\sigma)$. When $\bar{\beta}^{l}(\sigma) > \bar{\beta}^{FT-K}(\sigma)$, the equilibrium structure is as just described; (ml, sm) for $\beta \in \left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}^{l}(\sigma)\right)$, and (sl) for $\beta \in \left(\bar{\beta}^{l}(\sigma), 1\right)$ which is depicted in the middle pane of Figure 9.[23] However, once $\bar{\beta}_2(\sigma) < \bar{\beta}^{FT-K}(\sigma)$, then l prefers (sl) over (ml, sl, sm) even when the FT–K condition is satisfied. This is depicted in the bottom pane of Figure 9. Here, the benefit l foregoes by not becoming an insider with m and the hub is more than compensated by being able to permanently maintain insider rents with s. Therefore, (ml, sm) does not exist in equilibrium. Free trade is then attained when $\beta < \max\left(\bar{\beta}^{l}(\sigma), \bar{\beta}^{FT-I}_{s,l}(\sigma)\right)$.

The final complication which is not illustrated in Figure 9 is that when α_{ls} is sufficiently large, $\alpha_{ls} > 3-3\tau$, l prefers the status quo of no BAs over (sl). This affects the dynamic incentive structure because if l refuses to form a BA with s, then s can only form a BA with m which subsequently leads to (sm, ml). In this situation, by refusing to form a BA with either s or m, l can ensure the status quo of $(\emptyset)$ is the FDNE if s prefers$(\emptyset)$ over (sm, ml).

5 Equilibrium of the bilateralism and multilateralism games

To isolate the effect of BAs in a world where BAs and MFN agreements simultaneously exist, this section compares the equilibrium outcomes of two games. On the one hand, it considers the game where the option to form MFN agreements exists in addition to BAs — the bilateralism game. This is achieved by merely adding the MFN option to the analysis of the pure bilateralism game from Section 4. On the other hand, it considers the game where only MFN agreements exist — the multilateralism game.

Modeling of MFN agreements in the literature follows one of two approaches. Aghion et. al. (2007) allow a once and for all choice between global free trade and the status quo of no trade agreements. Saggi and Yildiz (2010) differ slightly by allowing all three countries or a pair of countries to form agreements

[23] It is easy to show that $\bar{\beta}^{l}(\sigma) > 1$ implies $\bar{\beta}^{FT-I}_{s,l}(\sigma) > \bar{\beta}^{FT-K}(\sigma)$.

that abide by the MFN principle, where the MFN tariff is assumed to be the member's joint optimal tariff. While the three country joint optimal tariff is zero in their model, and so the three country MFN agreement corresponds to global free trade, the two country joint optimal tariff is nonzero. In the model of this paper, the three country joint optimal tariff is also zero. Conversely, the two country joint optimal tariff is either the initial common tariff or zero. Thus, to avoid redundancy, it will be assumed that MFN agreements embody zero tariffs whether it is a two or three country MFN agreement. The three country MFN agreement is available both from the status quo situation of no agreements as well as from the situation where a single agreement exists whether it be MFN consistent or not.

Relative to the pure bilateralism game, adding the possibility of MFN agreements does not alter the analysis of Section 4 in any substantive way. Thus, the intuition of the commitment problem provided by the pure bilateralism game is crucial to understanding the effect of BAs in a world where MFN agreements also exist. The following proposition formalizes this idea.

Proposition 5.1. *Any FDNE of the pure bilateralism game remains an FDNE when MFN agreements are allowed. A unique FDNE of the pure bilateralism game remains the unique FDNE when MFN agreements are allowed, except for* (ml, sl, sm) *when* $\tau < \underline{\tau}$.

The intuition behind Proposition 5.1 is simple: rents earned along the equilibrium path make a direct move to free trade undesirable. When the unique FDNE is merely two insiders forming the sole BA, these insiders earn rents and protect them by blocking a direct move to free trade. When (ml, sm) is the unique FDNE, m earns rents as both an insider and the hub while l earns rents as an insider and benefits from being able to remain a spoke rather than move to free trade. When (ml, sl, sm) is the unique FDNE, there are some more subtle considerations. First, from the insider–outsider network, the smaller insider and the outsider would prefer to move directly to free trade. This obviously benefits the outsider, but also the smaller insider becaue it knows the larger insider will be the hub. Nevertheless, this move requires the consent of the larger insider who would forego the rents associated with being a hub. Since the outsider cannot commit to not form a BA with the larger insider, the larger insider will block a direct move to free trade to ensure it becomes the hub on the path to free trade. The second consideration is that, from the status quo of no agreements, m may actually benefit from the insider rent to the extent that it compensates for being discriminated against as a spoke. Indeed, there is only a very small section of the parameter space for which this is not true and (ml, sl, sm) is the unique FDNE equilibrium of the pure bilateralism game. To rule out this case, a very crude sufficient condition is $\tau < \underline{\tau} \approx .04$.

In the multilateralism game, when the only option for trade liberalization is MFN agreements, any country that is party to such an agreement could veto it. In the model of this paper, a two country MFN agreement is the worst possible outcome for the larger country in the agreement. Not only would this country prefer to be the third party country and free ride on the two country MFN agreement, the MFN access extended to the third party country is so large that the larger insider prefers the status quo of no agreements. When it comes to the three country MFN agreement, i.e. global free trade, the incentives of l are crucial. For s and m, absence of discrimination in l's market outweighs lost domestic profits and ensures they will not block global free trade. However, this is not always true for l. As the largest export market, it incurs the largest cost of preferential access via lost domestic profits while it gains the smallest benefit of preferential access via profits earned in its partner markets. For l to prefer global free trade over the status quo of no agreements, it must receive sufficient benefits from preferential access with m and s relative to the cost it incurs. The following proposition formalizes this discussion.

Proposition 5.2. *When only MFN agreements are allowed, moving to global free trade immediately is the unique FDNE iff l does not block global free trade over the status quo of no trade agreements. Otherwise, this status quo is the unique FDNE. l will block global free trade iff $\alpha_{ls} > \alpha_{ms} + 1 - 3\tau$.*

With an understanding of the equilibria under both the bilateralism game and the multilateralism game, the focus now turns to comparing the equilibrium outcomes under. Such a comparison falls into one of four cases. First, BAs are strong building blocs when global free trade can be attained under the bilateralism game but not under the multilateralism game. Second, BAs are weak building blocs when global free trade can be attained under both games. In the former case, BAs are both necessary and sufficient for attaining global free trade while they are only sufficient in the latter case. Third, BAs are strong stumbling blocs when global free trade is attained under the multilateralism game but cannot be attained under the bilateralism game. Here, the presence of BAs prevent attainment of global free trade. Finally, BAs are weak stumbling blocs when global free trade is not attained in either game.

Figure 10 shows the extent of each of these four cases when countries are similar and $\tau = \frac{1}{4}$. For the equilibrium of the multilateralism game, the band depicted by the $\alpha_{ms} = \alpha_{ls}$ and $\alpha_{ms} = \alpha_{ls} - 1 + 3\tau$ lines is critical. This band represents the set of $(\alpha_{ms}, \alpha_{ls})$ for which l will not block the three country MFN agreement leading to global free trade when only MFN agreements exist meaning the three country MFN agreement is the FDNE. To the right of this band is the set of $(\alpha_{ms}, \alpha_{ls})$ for which l will block the three country MFN agreement menaing there is no trade agreement in the FDNE. For the equilibrium of the bilateralism game, the line labelled $\bar{\beta} = \frac{1}{2}$ is critical. This line is a contour curve, chosen arbitrarily for illustration, showing the pairs of $(\alpha_{ms}, \alpha_{ls})$ such that, $\bar{\beta}_{m,l}^{FT-I}(\sigma) = \frac{1}{2}$. This contour curve was chosen arbitrarily for illustration. Because higher degrees of asymmetry reduce $\bar{\beta}_{m,l}^{FT-I}(\sigma)$, the contour curve is downward sloping and a higher contour curve indicates a lower $\bar{\beta}_{m,l}^{FT-I}(\sigma)$. As such, the upper contour set is the set of $(\alpha_{ms}, \alpha_{ls})$ such that $\bar{\beta}^{FT-I}_{m,l}(\sigma) < \frac{1}{2}$ meaning that m and l can overcome the commitment problem and remain insiders. Thus, by Proposition 5.1, the unique FDNE is (ml). Conversely, the lower contour set is the set of $(\alpha_{ms}, \alpha_{ls})$ such that m and l cannot overcome the commitment problem. In this case, a unique FDNE leads to free trade and, in the case of multiple FDNE, at most one FDNE does not lead to free trade (multiple FDNE are depicted by Θ^{BB} in Figure 10). Thus, in the equilibrium of the bilateralism game, BAs are stumbling blocs above the $\bar{\beta} = \frac{1}{2}$ contour curve while they are building blocs below the contour curve.

There is a key tension underlying the breakdown of the four cases depicted in Figure 10. On the one hand, the aggregate degree of asymmetry determines the ability of m and l to overcome the commitment problem; higher α_{ms} or α_{ls} increase their rents earned as insiders. On the other hand, the relative degree of asymmetry, $\frac{\alpha_{ls}}{\alpha_{ms}}$, drives whether l will block the three country MFN agreement leading directly to global free trade. If this ratio is sufficiently high, the cost of preferential access for l under free trade outweighs the benefits of preferential access it obtains. Thus, when aggregate asymmetry is sufficiently high, m and l can commit to remain insiders and whether BAs are strong or weak stumbling blocs depends on the relative degree of asymmetry. If it is not sufficiently biased, l will not block the MFN agreement moving the world to free trade and BAs are strong stumbling blocs; otherwise, BAs are only weak stumbling blocs. Conversely, when aggregate asymmetry is sufficiently low, m and l cannot commit to remain insiders. If the relative degree of asymmetry is sufficiently biased, l will block the three country MFN agreement meaning BAs are strong building blocs; otherwise, they are only weak building blocs.

Figure 10 also shows the effect of asymmetry on the strong building bloc–strong stumbling bloc issue. In particular, it shows that greater degrees of asymmetry weakly increase the extent to which BAs are strong and weak stumbling blocs and, eventually, decrease the extent to which BAs are strong and weak building blocs. The "weakly" and "eventually" caveats arise merely because the possibility of BAs being stumbling blocs depends on some minimum level of asymmetry. However, breakdown of the commitment problem between m and l causes a once off jump in the extent to which BAs are strong building blocs and a once off drop in the extent to which BAs are strong stumbling blocs. Thus, once again, understanding the equilibrium of the pure bilateralism game and the role of BAs as building blocs and stumbling blocs is crucial to understanding their role as strong building blocs and strong stumbling blocs.

Proposition 5.3. *Suppose a nontriviality condition holds,* $\beta > \bar{\beta}^{FT-I}_{m,l}(3-5\tau, 3-5\tau, \tau)$. *Then, breakdown of the commitment problem between l and m leads to a once off increase in the extent to which BAs are strong building blocs and a once off decrease in the extent to which BAs are strong stumbling blocs.*

Figure 11 extends Figure 10 to illustrate Proposition 5.3. The nontriviality condition just ensures that the discount factor is above that given by the $\bar{\alpha}_{ms} = \alpha_{ls}$ contour curve when countries are similar and

$\alpha_{ls} = 3 - 5\tau = 1.75$ in Figure 8. Once the commitment problem breaks down, the upper contour set of the $\bar{\beta}$ contour curves capture the set of $(\alpha_{ms}, \alpha_{ls})$ in which free trade is attained in an FDNE of the bilateralism game; the converse is true for the lower contour set. Thus, by Propositions 4.4 and 4.6, the upper contour set represents the set of $(\alpha_{ms}, \alpha_{ls})$ such that $\bar{\beta} < \bar{\beta}^{min}(\sigma)$ or, where relevant, $\bar{\beta} < \bar{\beta}^{l}(\sigma)$. That the upper contour set relates to the attainment of free trade is the opposite to the case when countries are similar. The reason follows the discussion after Proposition 4.6: a higher α_{ms} provides greater incentive for l to become the hub when the commitment problem breaks down but provides less incentive when overcoming the commitment problem is a possibility. Finally, the $\bar{\alpha}_{ms} = \alpha_{ls}$ contour curve represents the pairs of $(\alpha_{ms}, \alpha_{ls})$ for which the constraint that l and m are similar is strictly binding; the upper (lower) contour set satisfies (violates) this constraint.

Figure 11 dramatically illustrates Proposition 5.3 for the case of $\beta = \frac{1}{2}$. Here, conditional on a breakdown of the commitment problem, the $\bar{\beta} = \frac{1}{2}$ contour curve lies below the $\alpha_{ms} = \underline{\alpha}_{ms}$ contour curve. Thus, while BAs were either weak stumbling blocs or strong stumbling blocs for $\alpha_{ls} = 3 - 5\tau$ and $\beta = \frac{1}{2}$ because the commitment problem could be overcome for any α_{ms}, breakdown of the commitment problem reduces the restraint exercised by l in becoming the hub to the extent that BAs are strong or weak building blocs for any α_{ms}. Moreover, in Figure 11, BAs are largely strong building blocs because the relevant degrees of asymmetry are mostly biased meaning the benefit of preferential access that l gains in m's market is sufficiently small that it will block the three country MFN agreement. The multiplicity of equilibria, as indicated by Θ^{FT}, arises because the case of $\beta = \frac{1}{2}$ corresponds to the FDNE being in the set Θ from the bottom panes of Figure 9.

The extent to which breakdown of the commitment problem affects the equilibrium can be seen another way in Figure 12. Here, even though the discount factor is $\bar{\beta} = \frac{3}{4}$, which decreases the incentive for l to become the hub by increasing the importance of subsequent negative externalities, the extent to which BAs are strong building blocs has still increased dramatically because of the commitment problem breakdown. For $\bar{\beta} = \frac{3}{4}$, the role of BAs as weak stumbling blocs also arises when the cost of preferential access is sufficiently large relative to the benefits of preferential access as measured by $\frac{\alpha_{ls}}{\alpha_{ms}}$. In terms of Figure 9, Figure 12 falls under the case where $\bar{\beta}^{FT-K}(\sigma) > \bar{\beta}^{FT-I}_{s,l}(\sigma)$. The flatter segment of the $\bar{\beta} = \frac{3}{4}$ contour curve represents the case from the bottom pane of Figure 9 where $\bar{\beta}^{l}(\sigma) < \bar{\beta}^{FT-K}(\sigma)$ while the steeper segment represents the converse case from the middle pane of Figure 9. As shown in Figure 12, there are a few distinct FDNE in which BAs are weak stumbling blocs. Relative to Figure 11, this arises because the negative externalities are more important with the higher discount factor. The first such FDNE is (sl) which arises because of the increased ability and desirability for s and l to overcome their commitment problem. The second such FDNE is (ml, sm) which arises because l may prefer to let m be the hub rather than do so itself. Finally, the FDNE can consist of no trade agreements when $\frac{\alpha_{ls}}{\alpha_{ms}}$ is sufficiently large. Here, l prefers the status quo of no agreements and s cannot credibly commit to become an insider with m, meaning it cannot induce l to forego this status quo and become an insider.

Given the significant role of BAs as strong building blocs, which only arises because l could block the three country MFN agreement leading to global free trade, it is natural to wonder how each country is affected by the presence of BAs.

Proposition 5.4. *Suppose the FDNE of the bilateralism game is unique and BAs are strong building blocs. BAs increase the profits of m, while l must be sufficiently impatient and s must be sufficiently patient for their profits to rise.*

Since (ml, sl, sm) is the only unique FDNE in which free trade is attained, Proposition 5.4 largely revolves around Proposition 5.2 which says both s and m prefer free trade over the status quo of no agreements while l may not. Even though s and m are discriminated against on the path to free trade as an outsider and/or a spoke, there may be enough benefits on this path relative to the status quo of no agreements to outweigh such discrimination. For m, this is always so because of the rents earned as an insider. However, because of the discrimination faced as an outsider, s must be sufficiently patient. In contrast, despite the rents earned as an insider and a hub, l must be sufficiently impatient because of the possibility it may prefer the status quo of no agreements over global free trade. In Figure 12, each country experiences higher profits as a result of BAs being strong building blocs when l and s are nonsimilar, while this is rarely true when l and s are similar. The interesting implication is that BAs may not be Pareto improving under small degrees of asymmetry while they may indeed be Pareto improving under larger degrees of asymmetry.

Figure 12 also shows that large parts of the parameter space are not characterized by BAs being strong building blocs. The following proposition characterizes this case which revolves around rents earned or discrimination faced along the equilibrium path of the pure bilateralism game relative to the multilateralism game.

Proposition 5.5. *Suppose the FDNE is unique. When BAs are not strong building blocs, BAs increase the profits of l with one potential exception — the FDNE of (ml, sm). For BAs to increase the profits of s or m, they must be an insider.*

The reason l enjoys such an advantageous position is because as the largest export market it offers the greatest benefit of preferential access to its partners. As a result, l can form coalitions to ensure its preferred path arises in equilibrium. Proposition 5.5 is specific to unique FDNE because a multiplicity of equilibria arises when there is no coalition which can ensure its preferred path. Nevertheless, something similar to

Proposition 5.5 goes through when there are multiple FDNE because what drives Proposition 5.5 is the network positions that countries occupy.

When BAs are weak building blocs in the unique FDNE, the FDNE is (ml, sl, sm). Here, BAs benefit l because it is both an insider and the hub from which it earns rents it wouldn't under the MFN equilibrium of global free trade. While m also earns rents as an insider, it will be discriminated against as a spoke, so sufficient impatience is needed to ensure the insider rents dominate. Since the outsider is discriminated against as both an outsider and a spoke, the profits of s fall.

When countries are similar, Figure 12 shows BAs can be strong stumbling blocs with m and l remaining insiders. The rents they protect ensure higher profits relative to global free trade. However, s suffers from this discrimination. Although it does not happen in Figure 12, BAs could be strong stumbling blocs when the commitment problem breaks down. If so, (sl) or (ml, sm) could be the FDNE under bilateralism. In each case, l benefits from the rents as an insider and as a spoke. The latter follows because, here, l prefers being a spoke over global free trade. Additionally, m would benefit as an insider and the hub and s would benefit from the insider rents.

Finally, Figure 12 shows BAs can be weak stumbling blocs both when the commitment problem can and cannot be overcome. In the former case, m and l benefit from insider rents while s suffers from the discrimination it faces. In the latter case, the same analogy is true when s and l are insiders. BAs could also be weak stumbling blocs because of the hub–spoke network (ml, sm). Again, BAs benefit m because it is both an insider and the hub. However, while l benefits from being an insider and benefits by remaining a spoke as it avoids global free trade, it may still suffer overall because the profits as a spoke may be lower than the status quo of no agreements.

6 Conclusion

This paper makes two contributions to the current literature on preferential trade agreements. First, it identifies the importance of a commitment problem in understanding whether BAs are necessary for global free trade or necessarily lead away from global free trade when MFN agreements also exist. This commitment problem revolves around the desire of insiders to commit to not form another BA. Insiders recognize that by becoming the hub they subject themselves to negative externalities when the spokes subsequently form their own BA, and this motivates a desire to exercise mutual restraint in becoming the hub. The importance of the commitment problem emerges in a novel network framework with a new equilibrium concept, and this technical contribution represents the paper's second contribution. Primarily, the dynamic and farsighted aspects of the model allow emergence of the commitment problem. However, the new equilibrium concept ultimately arises because a coalitional simultaneous move game is embedded within the overall dynamic framework which, essentially, endogenizes the order in which countries negotiate BA formation.

A crucial aspect of the commitment problem is that the ability of insiders to overcome it depends intimately on market size asymmetry. A key implication is that asymmetry, through the commitment problem, has highly nonlinear effects on the equilibrium. While higher degrees of asymmetry initially increase the ability of insiders to overcome the commitment problem, which reduces the extent to which BAs lead to global free trade, there is a threshold level of asymmetry above which the commitment problem cannot be overcome. This causes a dramatic increase in the extent to which BAs lead to global free trade and, in turn, a dramatic increase (decrease) in the extent to which BAs are strong building (stumbling) blocs.

With the assumption of binding agreements, the model provides a rationale for the empirical observation of sequential BA negotiation: countries can form BAs on the equilibrium network path which exclude other countries. Countries exploit this opportunity because preferential access is more valuable when a country has sole preferential access, which leads to the existence of rents. That countries exploit sequential BA formation in response to the existence of rents also has two further implciations. First, even in the nonbargaining environment used here, the largest country often manipulates the network path to its advantage because it has the largest export market and can offer the greatest preferential access. Sometimes this manipulation comes at the expense of the other countries, yet it can also be consistent with BAs leading to a Pareto superior outcome relative to their absence. Second, even though answering the strong building bloc–strong stumbling bloc question depends on comparing the equilibrium outcome when only MFN agreements exist with the equilibrium outcome when BAs also exist, the latter is essentially equivalent to the equilibrium outcome when only BAs exist.

The cost of assuming countries are perfectly farsighted, in both the simultaneous and dynamic dimensions of the model, is that agreements are assumed to be binding. To assume nonbinding agreements leads into potential problems of cycles in sequential games, which are not observed in reality. While some reduced form justifications were given for assuming binding agreements, there is clear scope for further structural modeling.

Appendix

A One period network dependent profits

Let $\tilde{A}^2 = \left(A^l\right)^2 + \left(A^m\right)^2 + \left(A^s\right)^2$. Then $\pi_i^{I,j} = \frac{1}{16}\left[\left(A^i+\tau\right)^2 + \left(A^j+\tau\right)^2 + \left(A^k - 2\tau\right)^2\right]$ which reduces to $\frac{1}{16}\left[\tilde{A}^2 + 2\tau\left(A^i + A^j\right) - 4\tau A^k + 6\tau^2\right]$. Similarly, $\pi_i^H = \frac{1}{16}\left[\tilde{A}^2 + 2\tau\left(A^j + A^k\right) + 2\tau^2\right]$, $\pi_i^{FT} = \frac{1}{16}\tilde{A}^2$, $\pi_i^N = \frac{1}{16}\left[\tilde{A}^2 + 4\tau A^i - 4\tau\left(A^j + A^k\right) + 12\tau^2\right]$, $\pi_i^{K,j} = \frac{1}{16}\left[\tilde{A}^2 + 2\tau A^i - 6\tau A^j + 10\tau^2\right]$, $\pi_i^O = \frac{1}{16}\left[\tilde{A}^2 + 4\tau A^i - 6\tau\left(A^j + A^k\right) + 22\tau^2\right]$ The common tariff and nonprohibitive tariff assumptions are analytically convenient. They imply the difference between any pair of one period profits is independent of $\tilde{A}^2$. Then, the difference also has a common factor of τ. So, for example, $\pi_i^{I,j} - \pi_i^{FT} \propto A^i + A^j - 2A^k + 3\tau$. Since the proofs continually rely on linear combinations of differences in one period profits, this simple proportionality representation is extremely useful.

B Proofs

Proof of Lemma 2.1. Before beginning, the following lemma will be used throughout the proofs.

Lemma B.1. *The union of consistent sets is consistent.*

Proof. This is a known result. See Jackson (2008).■

For part a), suppose $\mathcal{C}(g) \subseteq \mathcal{T}(g)$ with $T(g) \in \mathcal{T}(g)$. Take some $\hat{T}(g) \in \mathcal{T}(g)$, $\hat{T}(g) \neq T(g)$. Since there are only three players, conditions i) and ii) imply that $T(g)$ is obtainable from any $\hat{T}(g)$ by S and there is no simultaneous farsighted improving path from $T(g)$. This has two implications. First, the deviation by S cannot be deterred. This implies $\hat{T}(g) \notin \mathcal{C}(g)$. Second, the simultaneous farsighted improving path from $\hat{T}(g)$ to $T(g)$ by S deters any obtainable deviation $\hat{T}(g)$ from $T(g)$. Thus, $T(g)$ is a consistent set and, by Lemma B.1, $T(g) = \mathcal{C}^*(g)$.

Now, part b). Consider a farsightedly stable M. Suppose that for each network on the equilibrium path, $T(g) = \mathcal{C}^*(g)$. By definition, M induces a unique $P(M, \emptyset)$ which is the unique FDNE.■

Proof of Proposition 4.1. All proofs of the FDNE will use backward induction to build up a farsightedly stable proposal map. By Lemmas 3.2 and 2.1, $\mathcal{C}^*(ij, ik) = (ij, ik, jk)$. Let $\mathcal{C}^*(ij, ik) \in M$. The remainder of the proof consists of two cases.

First, let $\beta < \bar{\beta}^{FT-I}(\sigma)$. Consider $\mathcal{C}(ij)$ and suppose $\mathcal{C}(ij) = \mathcal{T}(ij)$. $\beta < \bar{\beta}^{FT-I}(\sigma)$ implies each insider wants to become the hub and, because of symmetry and Lemma 3.2, the outsider wants to form a

BA yet is indifferent to its partner. This has two implications. First, because of the outsiders indifference, $\{(ij,ik),(ij,jk)\} \subseteq \mathcal{C}(ij)$. Second, there is no farsighted improving path that could deter the deviation from $T(ij)=(ij)$ to $\hat{T}(ij)=(ij,ik)$ or $\hat{T}(ij)=(ij,jk)$ which implies $(ij) \notin \mathcal{C}(ij)$. Thus, by Lemma B.1, $\mathcal{C}^*(ij)=\{(ij,ik),(ij,jk)\}$. Let $\mathcal{C}(ij) \in M$ for some $\mathcal{C}(ij) \in \mathcal{C}^*(ij)$. The following assumption will be used now and throughout the proofs.

Assumption 1. *When $\mathcal{C}^*(\emptyset)$ is not a singleton, each country believes each $T(g) \in \mathcal{C}^*(\emptyset)$ is equally likely to arise in equilibrium.*

Now consider $\mathcal{C}(\emptyset)$ and suppose $\mathcal{C}(\emptyset)=\mathcal{T}(\emptyset)$. Let $T(\emptyset)=(ij)$, $\hat{T}(\emptyset)=(\emptyset)$ and $S=ij$. Then, under Assumption 1, $P(M,(\emptyset)) \succ_S P\left(\hat{M},(\emptyset)\right)$ if $\frac{2}{3}f_1(\sigma)+\frac{1}{3}f_2(\sigma)>0$ where $f_1(\sigma)=\left(\pi_i^{I,j}+\frac{1}{2}\beta\left(\pi_i^H+\pi_i^{K,k}\right)+\frac{\beta^2}{1-\beta}\pi_i^{FT}\right)-\frac{1}{1-\beta}\pi_i^N$ and $f_2(\sigma)=\left(\pi_i^O+\beta\pi_i^{K,k}+\frac{\beta^2}{1-\beta}\pi_i^{FT}\right)-\frac{1}{1-\beta}\pi_i^N$. Algebra reveals $\frac{2}{3}f_1(\sigma)+\frac{1}{3}f_2(\sigma)>0$ holds. Again, because countries are symmetric, an insider is indifferent to its partner implying i)$\{(ij),(ik),(jk)\} \subseteq \mathcal{C}(\emptyset)$ and ii) there is no simultaneous farsighted improving path that could deter the deviation from $\hat{T}(\emptyset)=(\emptyset)$ to $T(\emptyset)=(ij)$ implying that $(\emptyset) \notin \mathcal{C}(\emptyset)$. Thus, by Lemma B.1, $\mathcal{C}^*(\emptyset)=\{(ij),(ik),(jk)\}$. Therefore, letting $\mathcal{C}(ij) \in M$ for some $\mathcal{C}(ij) \in \mathcal{C}^*(ij)$, the set of FDNE is the set of paths to free trade when $\beta < \bar{\beta}^{FT-I}(\sigma)$.

Now let $\beta \geq \bar{\beta}^{FT-I}(\sigma)$. Unlike earlier, $\beta \geq \bar{\beta}^{FT-I}(\sigma)$ implies both insiders prefer to remain insiders. Thus, by Lemma 2.1, $\mathcal{C}^*(ij)=(ij)$. Let $\mathcal{C}^*(ij) \in M$. For $\mathcal{C}^*(\emptyset)$, the same argument follows as for the case of $\beta < \bar{\beta}^{FT-I}(\sigma)$, except that $f_1(\sigma)=\frac{1}{1-\beta}\left(\pi_i^{I,j}-\pi_i^N\right)$ and $f_2(\sigma)=\frac{1}{1-\beta}\left(\pi_i^O-\pi_i^N\right)$. Therefore, the set of FDNE is the set of insider–outsider networks when $\beta \geq \bar{\beta}^{FT-I}(\sigma)$.■

Proof of Proposition 4.2. By Lemma 4.1, $\mathcal{C}^*(ij,jk)=(ij,ik,jk)$ so let $\mathcal{C}^*(ij,jk) \in M$. The following lemma will prove useful now and later.

Lemma B.2. *Suppose $\beta \in (0,1)$. Also suppose i and k as well as j and k are similar where $A^i > A^j$. Then, $\mathcal{C}^*(ij)=\mathcal{T}(ij)$ if $\beta \in \left[\bar{\beta}_{i,j}^{FT-I}(\sigma),\bar{\beta}_{j,i}^{FT-I}(\sigma)\right)$. Otherwise, $\mathcal{C}^*(ij)=(ij)$ for $\beta \geq \bar{\beta}_{j,i}^{FT-I}(\sigma)$ and $\mathcal{C}^*(ij)=(ij,ik)$ for $\beta < \bar{\beta}_{i,j}^{FT-I}(\sigma)$. Moreover, $\beta \geq \bar{\beta}_{i,j}^{FT-I}(\sigma)$ implies $\pi_i^{I,j}-\pi_i^{FT}=\pi_j^{I,i}-\pi_j^{FT}>0$.*

Proof. By Lemmas 4.1 and 2.1, $\mathcal{C}^*(ij,ik)=(ij,ik,jk)$. Letting $A^i > A^j$ implies, by simple algebra, $\bar{\beta}_{j,i}^{FT-I}(\sigma) > \bar{\beta}_{i,j}^{FT-I}(\sigma)$. Lemma 2.1 then implies $\mathcal{C}^*(ij)=(ij)$ for $\beta \geq \bar{\beta}_{j,i}^{FT-I}(\sigma)$, and using Lemmas 4.1, $\mathcal{C}^*(ij)=(ij,ik)$ for $\beta < \bar{\beta}_{i,j}^{FT-I}(\sigma)$.

For $\beta \in \left[\bar{\beta}_{i,j}^{FT-I}(\sigma),\bar{\beta}_{j,i}^{FT-I}(\sigma)\right)$, it must be shown that for each obtainable deviation from any $T(ij) \in \mathcal{T}(ij)$ by S, there is a simultaneous farsighted improving path that deters some $i \in S$. Let $T_0(ij) \rightarrow_{S_1} T_1(ij) \rightarrow \ldots \rightarrow_{S_K} T_K(ij)$ denote a simultaneous farsighted improving path. First, take $T(ij)=(ij)$. $\hat{T}(ij)=(ij,ik)$ is obtainable by ik, but $(ij,ik) \rightarrow_i (ij)$ deters i and k. $\hat{T}(ij)=(ij,jk)$ is obtainable by jk but $(ij,jk) \rightarrow_{ik} (ij,ik)$ deters j. Second, take $T(ij)=(ij,ik)$. $\hat{T}(ij)=(ij)$ is obtainable yet not preferred by k and obtainable by i yet $(ij) \rightarrow_{jk} (ij, jk)$ deters i. $\hat{T}(ij) = (ij, jk)$ is obtainable by jk but $(ij, jk) \rightarrow_{ik} (ij, ik)$ deters j. Third, take $T(ij) = (ij, jk)$. $\hat{T}(ij) = (ij)$ is obtainable but not preferred by j or k. $\hat{T}(ij) = (ij, ik)$ is obtainable by ik but $(ij, ik) \rightarrow_{i} (ij)$ deters k.

Finally, by equation (1), $\beta \geq \bar{\beta}^{FT-I}_{i,j}(\sigma)$ and $\beta < 1$ implies $\pi_i^{I,j} > \pi_i^{FT}$ and it is easily verified that $\pi_i^{I,j} - \pi_i^{FT} = \pi_j^{I,i} - \pi_j^{FT}$.■

The following remark will also prove useful going forward.

Remark B.1. $\pi_i^H > \pi_i^{I,j} > \pi_i^{I,k} > \pi_i^{K,k} > \pi_i^O$ where $A^j > A^k$

First, consider the case of $\beta \geq \bar{\beta}^{FT-I}_{m,l}(\sigma)$. In general, Lemma B.2 gives $\mathcal{C}^*(ij) = (ij)$ and here $\mathcal{C}^*(ml) = (ml)$. Let $\mathcal{C}^*(ij) \in M$. By definition of $\beta \geq \bar{\beta}^{FT-I}_{m,l}(\sigma)$ and Remark B.1, Lemma 2.1 implies $\mathcal{C}^*(\emptyset) = (ml)$. Letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (ml) for $\beta \geq \bar{\beta}^{FT-I}_{m,l}(\sigma)$.

Second, consider the case of $\beta \in \left[\bar{\beta}^{FT-I}_{l,m}(\sigma), \bar{\beta}^{FT-I}_{m,l}(\sigma)\right)$. By Lemma B.2, $\mathcal{C}^*(ml) = \mathcal{T}(ml)$ so let $\mathcal{C}^*(ml) \in M$. Simple algebra reveals $\bar{\beta}^{FT-I}_{m,l}(\sigma) < \bar{\beta}^{FT-I}_{m,s}(\sigma) < \bar{\beta}^{FT-I}_{s,m}(\sigma)$ implying $\mathcal{C}^*(sm) = (sm, ml)$ by Lemma B.2. But, $\bar{\beta}^{FT-I}_{l,s}(\sigma) \gtrless \bar{\beta}^{FT-I}_{m,l}(\sigma)$ implies, by Lemma B.2, $\mathcal{C}^*(sl) = \mathcal{T}(sl)$ or $\mathcal{C}^*(sl) = (sl, ml)$. Nevertheless, it will now be shown that, under Assumption 1, m and l prefer $T(\emptyset) = (ml)$ over any other $\hat{T}(\emptyset) \in \mathcal{T}(\emptyset)$. Thus, by Lemma 2.1, $\mathcal{C}^*(\emptyset) = (ml)$.

Let $\mathcal{C}^*(sm) \in M$ and $\mathcal{C}^*(sl) \in M$. Then, by Remark B.1, it must be shown that $P(M, \emptyset) \succ_{ml} P\left(\hat{M}, \emptyset\right)$ where $T(\emptyset) = (ml)$ and $\hat{T}(\emptyset) \in \{(sl), (sm), (\emptyset)\}$. For m and l with respect to, respectively, $\hat{T}(\emptyset) = (sl)$ and $\hat{T}(\emptyset) = (sm)$, this follows from $\pi_i^{I,j} > \pi_i^{I,s} > \pi_i^{FT} > \pi_i^O$. For m and l with respect to, respectively, $\hat{T}(\emptyset) = (sl)$ and $\hat{T}(\emptyset) = (sm)$, the following condition is necessary for $i = l$ and sufficient for $i = m$ under Assumption 1: $f_i(\sigma) = (1-\beta)\left(\pi_i^{I,j} - \pi_i^{I,s}\right) + \frac{1}{3}\beta(1-\beta)\left[\left(\pi_i^{I,j} - \pi_i^H\right) + \left(\pi_i^{K,s} - \pi_i^H\right)\right] + \frac{1}{3}\beta^2\left(\pi_i^{I,j} - \pi_i^{FT}\right) > 0$. This reduces to $f_i(\sigma) = \frac{1}{3}\left(\pi_i^{I,j} - \pi_i^{FT}\right) > 0$. Finally, letting $\hat{T}(\emptyset) = (\emptyset)$, $f_m(\sigma) > 0$ and $\beta < \bar{\beta}^{FT-I}_{m,s}(\sigma)$ imply $P(M, \emptyset) \succ_i P\left(\hat{M}, \emptyset\right)$ for $i = m$ while it is assumed to hold for $i = l$.[24] Therefore, by Lemma 2.1, $\mathcal{C}^*(\emptyset) = (ml)$. Letting $\mathcal{C}^*(\emptyset) \in M$, the set of FDNE is $\{(ml), (ml, sl, sm), (ml, sm, sl)\}$ for $\beta \in \left[\bar{\beta}^{FT-I}_{l,m}(\sigma), \bar{\beta}^{FT-I}_{m,l}(\sigma)\right)$.

Now consider the third, and last, case of $\beta < \bar{\beta}^{FT-I}_{l,m}(\sigma)$. Simple algebra shows $\beta < \bar{\beta}^{FT-I}_{l,m}(\sigma)$ implies $\beta \leq \bar{\beta}^{FT-I}_{i,j}(\sigma)$ for any i, j. Thus, by Lemma B.2, $\mathcal{C}^*(ij) = (ij, ik)$ where $A^i > A^j$ for any i, j. Letting $\mathcal{C}^*(ij) \in M$, free trade will be attained from any insider–outsider network. By Remark B.1, i) l prefers $T(\emptyset) = (ml)$ over $T(\emptyset) = (sl)$ over $T(\emptyset) = (sm)$, ii) with some additional simple algebra, s prefers $T(\emptyset) = (sl)$ over $T(\emptyset) = (sm)$ over $T(\emptyset) = (ml)$ and iii) m prefers $T(\emptyset) = (ml)$ over $T(\emptyset) = (sl)$. Then, letting $T(\emptyset) = (ml)$ and $\hat{T}(\emptyset) = (sm)$, $P(M, \emptyset) \succ_m P\left(\hat{M}, \emptyset\right)$ if $f_m(\sigma) = \left(\pi_m^{I,l} - \pi_m^{I,s}\right) + \frac{\beta}{1-\beta}\left(\pi_m^{K,s} - \pi_m^H\right) > 0$

[24] While $\beta < \bar{\beta}^{FT-I}_{m,s}(\sigma)$ implies that m prefers $T(\emptyset) = (sm)$ over $\hat{T}(\emptyset) = (\emptyset)$, $\pi_l^{FT} \gtrless \pi_l^N$ implies it is possible that l prefers $T(\emptyset) = (ml)$ over $T(\emptyset) = (\emptyset)$ even when $\beta < \bar{\beta}^{FT-I}_{l,m}(\sigma)$. Since the range of the parameter space for which this is occurs is very small (need $\tau \lessapprox .1$, $\alpha_{ls} \gtrapprox 2.6 - 4\tau$, $\alpha_{ms} \lessapprox 1.2 - .5\tau$), this case will be ignored.

which reduces to $\beta < \bar{\beta}^m(\sigma) \equiv \frac{\pi_m^{I,l} - \pi_m^{I,s}}{\pi_m^H - \pi_m^{K,s}}$. Thus, Lemma 2.1 implies $\mathcal{C}^*(\emptyset) = (ml)$ for $\beta < \bar{\beta}^m(\sigma)$ and, letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (ml, sl, sm). The following lemma will prove useful in considering the case of $\beta \in \left[\bar{\beta}^m(\sigma), \bar{\beta}_{l,m}^{FT-I}(\sigma)\right)$ which will complete the proof.

Lemma B.3. *Given M, suppose each country has a different first and second preference over $\mathcal{T}(\emptyset)$ where its first two preferences involve itself as an insider. In addition, if for each i, $P(M, \emptyset) \succ_i P\left(\hat{M}, \emptyset\right)$ where $T(\emptyset) = (\emptyset)$ and $\hat{T}(\emptyset) = (jk)$, then $\mathcal{C}^*(\emptyset) = \mathcal{T}(\emptyset)$. Otherwise, $\mathcal{C}^*(\emptyset) = \mathcal{T}(\emptyset) \setminus (\emptyset)$.*

Proof. Suppose $\mathcal{C}(\emptyset) = \mathcal{T}(\emptyset)$. Take $T(\emptyset) = (ij)$ as i's first preference and, to avoid the case of Lemma 2.1, j's second preference. Using the notation from the proof of Lemma B.2, $\hat{T}(\emptyset) = (ik)$ is obtainable by ik, but $(ik) \rightarrow_{ij} (ij)$ deters i and k. $\hat{T}(\emptyset) = (jk)$ is obtainable by jk, but $(jk) \rightarrow_{ik} (ik)$ deters j. $\hat{T}(\emptyset) = (\emptyset)$ is obtainable by i or j but $(\emptyset) \rightarrow_{jk} (jk)$ deters i and $(\emptyset) \rightarrow_{ik} (ik)$ deters j. Thus, $(ij) \in \mathcal{C}(\emptyset)$ and, by symmetry and Lemma B.1, $\mathcal{T}(\emptyset) \setminus (\emptyset) \subseteq \mathcal{C}(\emptyset)$.

Now suppose, for each i, $P(M, \emptyset) \succ_i P\left(\hat{M}, \emptyset\right)$ where $T(\emptyset) = (\emptyset)$ and $\hat{T}(\emptyset) = (jk)$. Then for each $\tilde{T}(\emptyset) = (ij)$ obtainable from $T(\emptyset)$ by $S = ij$, $(ij) \rightarrow_{jk} (jk)$ deters i. Thus, by Lemma B.1, $\mathcal{T}(\emptyset) = \mathcal{C}^*(\emptyset)$. Otherwise, for some i, $\tilde{T}(\emptyset) = (ij)$ is obtainable from $T(\emptyset) = (\emptyset)$ by $S = ij$ but no simultaneous farsighted improving path deters i. Thus, by Lemma B.1, $\mathcal{T}(\emptyset) \setminus (\emptyset) = \mathcal{C}^*(\emptyset)$.■

To apply Lemma B.3 for $\beta \in \left[\bar{\beta}^m(\sigma), \bar{\beta}_{l,m}^{FT-I}(\sigma)\right)$, it is sufficient to show that $P(M, \emptyset) \succ_i P\left(\hat{M}, \emptyset\right)$ for $i = s, m, l$ where $\hat{T}(\emptyset) = (\emptyset)$ and $T(\emptyset) = (sm)$ for $i = s$, $T(\emptyset) = (ml)$ for $i = m$, $T(\emptyset) = (sl)$ for $i = l$. For $i = s, m$, this can be shown to hold numerically for $\beta < \bar{\beta}_{m,l}^{FT-I}(\sigma)$ while for l it is assumed.[25] However, whether $\mathcal{C}^*(\emptyset) = \mathcal{T}(\emptyset)$ or $\mathcal{C}^*(\emptyset) = \mathcal{T}(\emptyset) \setminus (\emptyset)$ depends on σ. Thus the set of FDNE when $\beta \in \left[\bar{\beta}^m(\sigma), \bar{\beta}_{l,m}^{FT-I}(\sigma)\right)$ is $P^{FDNE} = \{(sm, ml, sl), (sl, ml, sm), (ml, sl, sm)\}$ or $P^{FDNE} \cup (\emptyset)$. Finally, it is useful to record the following remark the first part of which was just used and the second part of which follows from simple algebra.

Remark B.2. m prefers the dynamic path (ml, sl, sm) over $(\emptyset)$. This implies s prefers the dynamic path (ml, sl, sm) over $(\emptyset)$.■

Proof of Proposition 4.3. By Proposition 4.2, the commitment problem can be overcome when $\beta > \bar{\beta}_{m,l}^{FT-I}(\sigma)$ and this is when BAs are stumbling blocs. Thus, the proof rests on showing $\frac{\partial \bar{\beta}_{m,l}^{FT-I}(\sigma)}{\partial \alpha_{ms}} < 0$ and $\frac{\partial \bar{\beta}_{m,l}^{FT-I}(\sigma)}{\partial \alpha_{ls}} < 0$. This follows from inspection since $\bar{\beta}_{m,l}^{FT-I}(\sigma) \equiv \frac{\pi_m^H - \pi_m^{I,l}}{\pi_m^H - \pi_m^{FT}} = \frac{3 - \alpha_{ms} - 2\tau}{1 + \alpha_{ls} + \tau}$.■

Proof of Proposition 4.4. s and l are the only countries who are not similar. By Lemmas 4.1 and 2.1, $\mathcal{C}^*(ij, jk) = (ij, ik, jk)$ except for $(ij, jk) \in \{(sm, ml), (ml, sm)\}$ in which case $\mathcal{C}^*(ij, ik) = (ij, ik)$. In either case, let $\mathcal{C}^*(ij, jk) \in M$. Lemmas 3.1 and 2.1 imply $\mathcal{C}^*(sm) = (sm, ml)$. Lemma B.2 gives

[25] Again, this assumption is only violated for a very small section of the parameter space and thus will be ignored.

$\mathcal{C}^*(sl)$. By Lemma 2.1, using Lemmas 4.1 and 3.1, $\mathcal{C}^*(ml) = (ml, sl)$ if $\beta < \bar{\beta}^{FT-K}(\sigma)$. It is trivial that $\mathcal{C}^*(ml) = (ml, sm)$ for $\beta \geq \bar{\beta}^{FT-K}(\sigma)$. In any case, let $\mathcal{C}^*(ij) \in M$.

$\mathcal{C}^*(\emptyset)$ remains to be determined. First, consider the critical values $\bar{\beta}^s(\sigma)$ and $\bar{\beta}^m(\sigma)$. Suppose $P(M, \emptyset) \succ_i P\left(\hat{M}, \emptyset\right)$ for $i = s, m$ where $T(\emptyset) = (ml)$ and $\hat{T}(\emptyset) = (sm)$. This requires $\triangle_s(\beta, \sigma) = (1-\beta)\left(\pi_s^O - \pi_s^{I,m}\right) + \beta(1-\beta)\left(\pi_s^{K,m} - \pi_s^{K,l}\right) + \beta^2\left(\pi_s^{FT} - \pi_s^{K,l}\right) > 0$ and $\triangle_m(\beta, \sigma) = (1-\beta)\left(\pi_m^{I,l} - \pi_m^{I,s}\right) + \beta(1-\beta)\left(\pi_m^{K,s} - \pi_m^H\right) + \beta^2\left(\pi_m^{FT} - \pi_m^H\right) > 0$. Implicitly, these reduce to $\beta > \bar{\beta}^s(\sigma)$ and $\beta < \bar{\beta}^m(\sigma)$ respectively. The following assumption will be used as a baseline in proceeding. Violations of this assumption occur in a small section of the parameter space and do not substantively affect the analysis; the implications of such violations will be noted.

Assumption 2. *Assume that i) $\mathcal{C}^*(sl) = (sl)$ and $\mathcal{C}^*(sl) \in M$ and ii) $P(M_0, \emptyset) \succ_l P(M_2, \emptyset)$ and $P(M_1, \emptyset) \succ_l P(M_2, \emptyset)$ where $T_0(\emptyset) = (ml)$, $T_1(\emptyset) = (sl)$ and $T_2(\emptyset) = (\emptyset)$.*

Suppose $\beta < \min\left(\bar{\beta}^m(\sigma), \bar{\beta}^{min}(\sigma)\right)$. Remarks B.1 and B.2 imply m's first preference is $T(\emptyset) = (ml)$. Independent of $\mathcal{C}^*(sl)$, simple algebra shows l prefers $T(\emptyset) = (sl)$ over $\hat{T}(\emptyset) = (sm)$. Remark B.1 then implies l's first and second preferences are, respectively, $T(\emptyset) = (ml)$ and $T(\emptyset) = (sl)$. Thus, Lemma 2.1 implies $\mathcal{C}^*(\emptyset) = (ml)$ and, letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (ml, sl, sm).

Under Assumption 2, the following lemma will prove $\mathcal{C}^*(\emptyset) = (ml)$ for $\beta \in \left(\bar{\beta}^s(\sigma), \bar{\beta}^{min}(\sigma)\right)$. Essentially, it provides conditions under which countries i and j can "force" k to form a link with i.

Lemma B.4. *Suppose there are three players. Given $\mathcal{T}(g)$ and some M, suppose that i) $T_0(g)$ is i's first preference and j's second preference and vice versa for $T_1(g)$ where $T_1(g) = g$ or $T_1(g) = (g, ij)$ and ii) k prefers $T_0(g)$ over $T_1(g)$ and $T_0(g)$ is obtainable from $T_1(g)$ by $S_0 = ik$. Then, $\mathcal{C}^*(g) = T_0(g)$.*

Proof. Suppose $\mathcal{C}(g) = \mathcal{T}(g)$. To avoid Lemma 2.1, assume that $P(M_2, g) \succ_k P(M_0, g)$ for some $T_2(g) \in \mathcal{T}(g)$. First, consider any $T_2(g) \in \mathcal{T}(g) \setminus \{T_0(g), T_1(g)\}$. By conditions i) and ii), the deviation to $T_1(g)$ by $S_1 = ij$ is obtainable and the only simultaneous farsighted improving path is $T_1(g) \rightarrow_{ik} T_0(g)$ which does not deter S_1. Thus, $T_2(g) \notin \mathcal{C}(g)$. Second, consider $T_1(g)$. By conditions i) and ii), $T_0(g)$ is obtainable from $T_1(g)$ by $S_0 = ik$ and preferred by S_0. Moreover, there is no simultaneous farsighted improving path from $T_0(g)$. Thus, $T_1(g) \notin \mathcal{C}(g)$. Finally, consider $T_0(g)$. Take any obtainable $\hat{T}(g)$ by $\hat{S}$. If $\hat{T}(g) = T_1(g)$, condition i) implies $T_1(g) \rightarrow_{ik} T_0(g)$ deters $\hat{S}$. Otherwise, $\hat{T}(g) \rightarrow_{ij} T_1(g) \rightarrow_{ik} T_0(g)$ deters $\hat{S}$ by conditions i) and ii). Therefore, by Lemma B.1, $\mathcal{C}^*(g) = T_0(g)$.■

Under Assumption 2, Lemma B.4 can be applied for $\beta \in \left(\bar{\beta}^s(\sigma), \bar{\beta}^{min}(\sigma)\right)$ because, with respect to $\mathcal{T}(\emptyset)$, three things are true given M. First, the first and second preferences of s are $T_1(\emptyset) = (sl)$ and $T_0(\emptyset) = (ml)$ respectively. This follows from Remark B.1, simple algebra and the fact that it can be shown numerically that s prefers $T_2(\emptyset) = (sm)$ over $T_3(\emptyset) = (\emptyset)$ for $\beta < \bar{\beta}^{min}(\sigma)$. Second, by Remark B.1 as well as simple algebra which shows l prefers $T_1(\emptyset) = (sl)$ over $T_3(\emptyset) = (sm)$, l's first and second preferences are, respectively, $T_0(\emptyset) = (ml)$ and $T_1(\emptyset) = (sl)$. Third, by Remark B.1, m prefers $T_0(\emptyset) = (ml)$ over $T_1(\emptyset) = (sl)$. Therefore, $\mathcal{C}^*(\emptyset) = (ml)$ for $\beta \in \left(\bar{\beta}^s(\sigma), \bar{\beta}^{min}(\sigma)\right)$ and, letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (ml, sl, sm).

The use of Lemma 2.1 in the case of $\beta < min\left(\bar{\beta}^m(\sigma), \bar{\beta}^{min}(\sigma)\right)$ and Lemma B.4 in the case of $\beta \in \left(\bar{\beta}^s(\sigma), \bar{\beta}^{min}(\sigma)\right)$ are sensitive to Assumption 2 ii). When l prefers $T(\emptyset) = (\emptyset)$ over $\hat{T}(\emptyset) = (ml)$ there may be multiple equilibria stemming from $\mathcal{C}^*(\emptyset)$ not being a singleton. Nevertheless (ml, sl, sm) is an FDNE except if Assumption 2 is violated and the first preference of both l and s is $T(\emptyset) = (sl)$. In this case, Lemma 2.1 implies $\mathcal{C}^*(\emptyset) = (sl)$ and the set of FDNE are $\{(sl), (sl, ml, sm), (sl, sm, ml)\}$.

Under Assumption 2, and given what has been shown already, it follows that Lemma B.3 can be applied to determine the set $\mathcal{C}^*(\emptyset)$ for $\beta \in \left(\bar{\beta}^m(\sigma), \bar{\beta}^s(\sigma)\right)$. Thus, free trade results in the FDNE conditional on $T(\emptyset) = (sl)$ or $T(\emptyset) = (ml)$. Again the use of Lemma B.3 is somewhat sensitive to Assumption 2. When l prefers $T(\emptyset) = (\emptyset)$ over $\hat{T}(\emptyset) = (ml)$, $\mathcal{C}^*(\emptyset) = \mathcal{T}^*(\emptyset)$ or $\mathcal{C}^*(\emptyset) = \mathcal{T}^*(\emptyset) \setminus (\emptyset)$ except when Assumption 2 i) is violated. Here, it is possible for $\mathcal{C}^*(\emptyset) = (sl)$ in which case the set of FDNE is, again, $\{(sl), (sl, ml, sm), (sl, sm, ml)\}$.■

Proof of Proposition 4.5. Let $l > s$ denote that l and s are not similar. Since $\frac{\partial \bar{\beta}^{FT-I}_{m,l}(\sigma)}{\partial \alpha_{ms}} = \frac{3 - \bar{\beta}^{FT-I}_{m,l}(\sigma)}{\pi^H_m - \pi^{I,l}_m} > 0$, let $\bar{\alpha}_{ms} = 1$ upon which $\bar{\beta}^{FT-I}_{m,l}(\bar{\alpha}_{ms}, \bar{\alpha}_{ls}, \tau) = \frac{1}{2}$. It needs to now be shown that $\bar{\beta}^{min}(\hat{\sigma}) \geq \frac{1}{2}$ for any $\hat{\sigma}$ satisfying $l \approx m$, $m \approx s$ and $l > s$, i.e $\alpha_{ls} \in [3 - 5\tau, 3 - 2\tau)$ and $\alpha_{ms} \in \left[1 + \frac{6-c}{3}\tau, 3 - 5\tau\right)$ where the parametrization $\alpha_{ls} = 3 - c\tau$ for $c \in (2, 5]$ is being used. First, consider $\bar{\beta}^{FT-K}(\sigma) = \frac{\pi^H_l - \pi^{K,s}_l}{\pi^H_l - \pi^{FT}_l} = \frac{-\alpha_{ls} + 4 + \alpha_{ms} - 4\tau}{\alpha_{ms} + 1 + \tau} \in (0, 1]$. This is minimized for $\tilde{\sigma} = \left(1 + \frac{1}{3}\tau, 3 - 2\tau, \tau\right)$ and $\tau = \frac{1}{3}$ implying $\bar{\beta}^{FT-K}(\tilde{\sigma}) = \frac{13}{22} > \frac{1}{2}$ as required. Second, consider $\bar{\beta}^{FT-I}_{s,l}(\sigma) = \frac{\pi^H_s - \pi^{I,L}_s}{\pi^H_s - \pi^{FT}_s} = \frac{-1 + 3\alpha_{ms} - 2\tau}{\alpha_{ms} + \alpha_{ls} + \tau}$. To ensure $l \approx m$, this can be written as $\bar{\beta}^{FT-I}_{s,l}(\sigma) = \frac{-1 + 3\left(1 + \frac{6-c}{3}\tau\right) - 2\tau}{\left(1 + \frac{6-c}{3}\tau\right) + (3 - c\tau) + \tau} = \frac{2 + 4\tau - c\tau}{4 + 3\tau - \frac{4}{3}\tau}$ where $\frac{\partial \bar{\beta}^{FT-I}_{s,l}(\sigma)}{\partial c} \gtrless 0$ as $\bar{\beta}^{FT-I}_{s,l}(\sigma) \gtrless \frac{3}{4}$. Simple algebra shows $\bar{\beta}^{FT-I}_{s,l}(\sigma) = \frac{2 + 4\tau - c\tau}{4 + 3\tau - \frac{4}{3}\tau} < \frac{3}{4}$ so $\frac{\partial \bar{\beta}^{FT-I}_{s,l}(\sigma)}{\partial c} < 0$. Thus, subject to $l \approx m$, $\bar{\beta}^{FT-I}_{s,l}(\sigma)$ is minimized for $c = 5$; i.e. for $\alpha_{ls} = \bar{\alpha}_{ls}$. Therefore, since $\bar{\beta}^{FT-I}_{s,l}(\sigma) > \bar{\beta}^{FT-I}_{m,l}(\sigma)$, it follows that $\bar{\beta}^{FT-I}_{s,l}(\hat{\sigma}) \geq \bar{\beta}^{FT-I}_{s,l}(\bar{\alpha}_{ms}, \bar{\alpha}_{ls}, \tau) > \bar{\beta}^{FT-I}_{m,l}(\bar{\alpha}_{ms}, \bar{\alpha}_{ls}, \tau) = \frac{1}{2}$ for any $\hat{\sigma}$ as required.

For the second part of the proposition, direct inspection reveals $\frac{\partial \bar{\beta}^{FT-I}_{s,l}(\sigma)}{\partial \alpha_{ls}} < 0$ and $\frac{\partial \bar{\beta}^{FT-K}(\sigma)}{\partial \alpha_{ls}} > 0$. Additionally, $\frac{\partial \bar{\beta}^{FT-K}(\sigma)}{\partial \alpha_{ms}} = \frac{1 - \bar{\beta}^{FT-K}(\sigma)}{\pi^H_l - \pi^{FT}_l} > 0$ because, by construction, $l > s$ implies $\bar{\beta}^{FT-K}(\sigma) < 1$. Finally, $\frac{\partial \bar{\beta}^{FT-I}_{s,l}(\sigma)}{\partial \alpha_{ms}} = \frac{3 - \bar{\beta}^{FT-I}_{s,l}(\sigma)}{\pi^H_s - \pi^{I,l}_s} > 0$ for $\bar{\beta}^{FT-I}_{s,l}(\sigma) < 1$. ■

Proof of Proposition 4.6. $\mathcal{C}^*(ij, jk)$ and $\mathcal{C}^*(ij)$ are as described in the proof of Proposition 4.4, noting that $\beta > \bar{\beta}^{min}(\sigma)$. In any case, let $\mathcal{C}^*(ij, jk) \in M$ and $\mathcal{C}^*(ij) \in M$. Yet to be determined is $\mathcal{C}^*(\emptyset)$.

For now, assume $\alpha_{ls} < 3 - 3\tau$ so that $\pi^{I,s}_l > \pi^N_l$. First, consider the case of $\bar{\beta}^{min}(\sigma) = \bar{\beta}^{FT-I}_{s,l}(\sigma)$. To begin, some critical values will be determined. Define $T_{ij}(\emptyset) = (ij)$ and $T_{\emptyset}(\emptyset) = \emptyset$ so that, for example, M_{ij} denotes that the proposal map contains $T_{ij}(\emptyset)$. If $\beta > \bar{\beta}^{FT-K}(\sigma)$, $P(M_{sl}, \emptyset) \succ_l P(M_{ml}, \emptyset)$ requires $(1-\beta)\left(\pi_l^{I,s} - \pi_l^{I,m}\right) + \beta\left(\pi_l^{I,s} - \pi_l^{K,s}\right) > 0$ which reduces to $\bar{\beta}^l(\sigma) > \frac{\pi_l^{I,m} - \pi_l^{I,s}}{\pi_l^{I,m} - \pi_l^{K,s}}$. However, if $\beta < \bar{\beta}^{FT-K}(\sigma)$ then $P(M_{sl}, \emptyset) \succ_l P(M_{ml}, \emptyset)$ requires $(1-\beta)\left(\pi_l^{I,s} - \pi_l^{I,m}\right) + \beta(1-\beta)\left(\pi_l^{I,s} - \pi_l^{H}\right) + \beta^2\left(\pi_l^{I,s} - \pi_l^{FT}\right) > 0$ which implicitly reduces to $\beta > \bar{\beta}^l(\sigma)$. Finally, note that i) $T_{sl}(\emptyset)$ is the first preference of s by Lemma B.2, Remark B.1, $\pi_s^{FT} > \pi_s^{K,j}$ and $\pi_l^{I,s} > \pi_l^{N}$, ii) $T_{ml}(\emptyset)$ is the second preference of s when $\beta > \bar{\beta}^s(\sigma)$ and iii) $T_{sl}(\emptyset)$ or $T_{ml}(\emptyset)$ is l's first preference, and the other its second, as determined by $\bar{\beta}^l(\sigma) \lesseqgtr \bar{\beta}^{FT-K}(\sigma)$ given Assumption 2 and Remark B.1.

Thus, by Lemma 2.1, $\mathcal{C}^*(\emptyset) = (sl)$ for $\beta \in \left(\max\left\{\bar{\beta}_{s,l}^{FT-I}(\sigma), \bar{\beta}^l(\sigma)\right\}, 1\right)$ and, by letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (sl). For $\beta \in \left(\max\left\{\bar{\beta}_{s,l}^{FT-I}(\sigma), \bar{\beta}^s(\sigma)\right\}, \bar{\beta}^l(\sigma)\right)$, $\mathcal{C}^*(\emptyset) = (ml)$ by Lemma 2.1 when $\bar{\beta}^l(\sigma) > \bar{\beta}^{FT-K}(\sigma)$ and by Lemma B.4 when $\bar{\beta}^l(\sigma) < \bar{\beta}^{FT-K}(\sigma)$ and $\beta > \bar{\beta}^s(\sigma)$. Then, letting $\mathcal{C}^*(\emptyset) \in M$, (ml, sm) is unique for $\beta > \bar{\beta}^{FT-K}(\emptyset)$ and (ml, sm, sl) is unique for $\beta \leq \bar{\beta}^{FT-K}(\emptyset)$ as long as $\bar{\beta}_{s,l}^{FT-I}(\sigma) > \bar{\beta}^s(\sigma)$. When this last condition does not hold, the analysis from Proposition 4.4 for $\beta \in \left(\bar{\beta}^m(\sigma), \bar{\beta}^s(\sigma)\right)$ applies.

Now, consider the second case of $\bar{\beta}^{min}(\sigma) = \bar{\beta}^{FT-K}(\sigma)$. Given what has been shown already, two implications follow from Lemma 2.1. First, $\mathcal{C}^*(\emptyset) = (ml)$ for $\beta \in \left(\bar{\beta}^{FT-K}(\sigma), \bar{\beta}^l(\sigma)\right)$ and, by letting $\mathcal{C}^*(\emptyset) \in M$, (ml, sm) is unique. Second, $\mathcal{C}^*(\emptyset) = (sl)$ for $\beta \in \left(\bar{\beta}^l(\sigma), 1\right)$ and, by letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is (sl).

Now, assume $\alpha_{ls} > 3 - 3\tau$ so that $P(M_{\emptyset}, \emptyset) \succ_l P(M_{sl}, \emptyset)$ and note $P(M_{sm}, \emptyset) \succ_s P(M_{\emptyset}, \emptyset)$ reduces to $\beta > \bar{\beta}_1(\sigma) = \frac{\pi_s^{I,m} - \pi_s^{N}}{\pi_s^{I,m} - \pi_s^{K,l}}$. Since it can be shown that $\bar{\beta}_1(\sigma) > \max\left(\bar{\beta}^{min}(\sigma), \bar{\beta}^l(\sigma)\right)$, then $\beta > \bar{\beta}_1(\sigma)$ implies $T(\emptyset) = (\emptyset)$ is l's first preference with respect to $\mathcal{T}(\emptyset) = (\emptyset)$. $\beta > \bar{\beta}_1(\sigma)$ also implies that, respectively, $T(\emptyset) = (sl)$ and $T(\emptyset) = (\emptyset)$ are s's first and second preference. It is then simple to show that, regardless of whether $T(\emptyset) = (sl)$ is l's second or third preference, Lemma B.4 can be applied. Thus, $\mathcal{C}^*(\emptyset) = (\emptyset)$ for $\beta > \bar{\beta}_1(\sigma)$ and, letting $\mathcal{C}^*(\emptyset) \in M$, the unique FDNE is $(\emptyset)$. For $\bar{\beta}_1(\sigma) < \max\left(\bar{\beta}^{min}(\sigma), \bar{\beta}^l(\sigma)\right)$ it is simple to show that the same results apply as for the case of $\alpha_{ls} < 3 - 3\tau$ with one exception — when $\mathcal{C}^*(\emptyset) = (sl)$ for $\alpha_{ls} < 3 - 3\tau$, now $\mathcal{C}^*(\emptyset) = \mathcal{T}(\emptyset)$.■

Proof of Proposition 5.1. The largest consistent set under the bilateralism and pure bilateralism games will be denoted $\mathcal{C}^{*MFN}(g)$ and $\mathcal{C}^*(g)$. $\mathcal{T}^{MFN}(g)$ and $\mathcal{T}(g)$ are similarly denoted. For the second part of the proposition, it needs to be shown that $\mathcal{C}^*(g) \subseteq \mathcal{C}^{*MFN}(g)\, \forall g \in \mathcal{G}$. Thus, for each $T(g) \in \mathcal{C}^*(g)$, obtainable deviations to each $\hat{T}(g) \in \mathcal{T}^{MFN}(g) \setminus \mathcal{T}(g)$ must be deterred. The first part of the proposition will be proved along the way with the relevant FDNE being $(\emptyset)$, (sl), (ml), (ml, sm) and (ml, sl, sm). Before beginning, note that Lemma 2.1 still applies when MFN agreements are allowed since it merely depends upon $T(g)$ being obtainable from any $\hat{T}(g)$ by S which remains true.

First, consider $\mathcal{C}^*(ij,jk)$. $\mathcal{T}(ij,jk) = \mathcal{T}^{MFN}(ij,jk)$ implies $\mathcal{C}^*(ij,jk) = \mathcal{C}^{*MFN}(ij,jk)$. Second, consider $\mathcal{C}^*(ij)$ noting that $\mathcal{T}^{MFN}(ij) = \mathcal{T}(ij) \cup (ijk)^{MFN}$. For $(ij) \in \mathcal{C}^*(ij)$, Lemma B.2 implies the deviation to $T(ij) = (ijk)^{MFN}$ by $S = ijk$ is not preferred by the larger insider i. Thus, $(ijk)^{MFN} \rightarrow_i (ij)$ deters S. For $(ij,ik) \in \mathcal{C}^*(ij)$, the deviation to $T(ij) = (ijk)^{MFN}$ by S is not preferred by the hub because $\pi_i^H > \pi_i^{FT}$. Thus, $(ijk)^{MFN} \rightarrow_i (ij) \rightarrow_{ik} (ij,ik)$ deters S. Therefore, $\mathcal{C}^*(ij) \subseteq \mathcal{C}^{*MFN}(ij)$. Moreover, by Lemma B.2, Lemma 2.1 implies $\mathcal{C}^*(ij) = \mathcal{C}^{*MFN}(ij)$ when $\mathcal{C}^*(ij) = (ij)$. It is also true that $\mathcal{C}^*(ij) = \mathcal{C}^{*MFN}(ij)$ when $\mathcal{C}^*(ij) = (ij,ik)$. This follows from three facts. First, the same logic by which $T(ij) \notin \mathcal{C}^*(ij)$ for $T(ij) \neq (ij,ik)$ still applies for $T(ij) \notin \mathcal{C}^{*MFN}(ij)$. Second, the obtainable deviation from $T(ij) = (ij,ik)$ to $\hat{T}(ij) = (ijk)^{MFN}$ by S is deterred by $(ijk)^{MFN} \rightarrow_i (ij) \rightarrow_{ik} (ij,ik)$. Third, take $T(ij) = (ijk)^{MFN}$ and consider the obtainable deviation by i to $T(ij) = (ij)$. The only simultaneous farsighted improving path is $(ij) \rightarrow_{ik} (ij,ik)$ which does not deter i and thus implies $(ijk)^{MFN} \notin \mathcal{C}^{*MFN}(ij)$.

Now consider $\mathcal{C}^*(\emptyset)$ noting that $\mathcal{T}^{MFN}(\emptyset) = \mathcal{T}(ij) \cup \left\{(ij)^{MFN}, (ik)^{MFN}, (jk)^{MFN}, (ijk)^{MFN}\right\}$ and that $\mathcal{C}^{*MFN}\left((ij)^{MFN}\right) = (ij)^{MFN}$ because ${}^M\pi_i^O > \pi_i^{FT}$ where the superscript M indicates the presence of an MFN agreement. Also note the following. First, under Assumption 2, the proofs of Propositions 4.2, 4.4 and 4.6 showed that when $\mathcal{C}^*(\emptyset)$ is not a singleton, (ml,sm,sl) is an FDNE. In these cases, by construction of $\bar{\beta}^m(\sigma)$ and $\bar{\beta}^s(\sigma)$, $T(\emptyset) = (ml)$ is preferred over $T(\emptyset) = (sm)$ by m and s. Second, by Remark B.1, $T(\emptyset) = (sl)$ is also preferred over $T(\emptyset) = (ml)$ by s. Third, by Remark B.2 and Assumption 2, m and l prefer $T(\emptyset) = (ml)$ over $\hat{T}(\emptyset) = (\emptyset)$. Fourth, $\pi_i^N > {}^M\pi_i^{I,j}$ for $A^i > A^j$. Given all of this, define the following simultaneous farsighted improving paths. Let P_l^F be $g^{MFN} \rightarrow_l (\emptyset) \rightarrow_{ml} (ml)$ where $g^{MFN} \in \left\{(sl)^{MFN}, (ml)^{MFN}, (sml)^{MFN}\right\}$. Let P_m^F be $(sm)^{MFN} \rightarrow_m (\emptyset) \rightarrow_{ml} (ml)$. Then, as relevant, and with one exception, P_l^F or P_m^F deters the obtainable deviation from $T(\emptyset) \in \mathcal{C}^*(\emptyset)$ to $\hat{T}(\emptyset) = g^{MFN}$ by S. The exception is for the obtainable deviation from $T(\emptyset) = (sl)$ to $\hat{T}(\emptyset) = (ml)^{MFN}$ by $S = ml$. Then, $(ml)^{MFN} \rightarrow_l (\emptyset) \rightarrow_{sl} (sl)$ deters S. Therefore, $\mathcal{C}^*(\emptyset) \subseteq \mathcal{C}^{*MFN}(\emptyset)$.

The insiders, if they are to remain so, prefer $(ij) = \mathcal{C}^*(\emptyset)$ over any MFN agreement involving themselves because of Lemma B.2 and $\pi_i^{I,j} > {}^M\pi_i^{I,j}$. Thus, although Lemma 2.1 does not strictly apply because ${}^M\pi_i^O > \pi_i^{I,j}$, the fact that $T(\emptyset) = (jk)^{MFN}$ is not obtainable by i makes it trivial that $\mathcal{C}^*(\emptyset) = \mathcal{C}^{*MFN}(\emptyset)$ when $\mathcal{C}^*(\emptyset) = \mathcal{C}^*(ij) = (ij)$. Under Assumption 2, the same is true for $\mathcal{C}^*(\emptyset) = (ml)$ when the unique FDNE of the pure bilateralism game is (ml,sm) or (ml,sl,sm) when $\tau > \underline{\tau}$. In the case of (ml,sl,sm) this follows because m and l prefer the dynamic path (ml,sl,sm) over $(sml)^{MFN}$. While for l this is true by $\pi_l^H > \pi_l^{I,m} > \pi_l^{FT}$, it is true for m by definition of $\underline{\tau}$.

While Assumption 2 has been used repeatedly in this proof, it is not essential and it is straightforward, although tedious, to work through the cases where Assumption 2 is violated.■

Proof of Proposition 5.2. Let the superscript M with respect to π indicate profits under the relevant

MFN agreement noting that ${}^{M}\pi_i^O > \pi_i^{FT} > {}^{M}\pi_i^{I,j}$ and, where $A^i > A^j$, $\pi_i^N > {}^{M}\pi_i^{I,j}$. Then, by Lemma 2.1, $\mathcal{C}^{*MFN}\left((ij)^{MFN}\right) = (ij)^{MFN}$ where $\mathcal{T}\left((ij)^{MFN}\right) = \left\{(ij)^{MFN}, (ijk)^{MFN}\right\}$. Now consider $\mathcal{C}^{*MFN}(\emptyset)$ noting that $\mathcal{T}(\emptyset) = \left\{(ij)^{MFN}, (ik)^{MFN}, (jk)^{MFN}, (ijk)^{MFN}\right\}$. Simple algebra shows that $\pi_i^{FT} > \pi_i^N$ for $i = s, m$ while it is only true for $i = l$ if $\alpha_{ls} < \alpha_{ms} + 1 - 3\tau$. Thus, there are two cases to consider. Initially, suppose $\alpha_{ls} < \alpha_{ms} + 1 - 3\tau$. Then each country prefers $T(\emptyset) = (ijk)^{MFN}$ over $\hat{T}(\emptyset) = (\emptyset)$. However, Lemma 2.1 cannot be applied because ${}^{M}\pi_i^O > \pi_i^{FT}$. Nevertheless, it is trivial to show that $\mathcal{C}^{*MFN}(\emptyset) = (ijk)^{MFN}$ because $(jk)^{MFN}$ is not obtainable by i. Thus, $(ijk)^{MFN}$ is the unique FDNE.

Now suppose $\alpha_{ls} \geq \alpha_{ms} + 1 - 3\tau$. Take any deviation from $T(\emptyset) = (\emptyset)$ to $\hat{T}(\emptyset) = g^{MFN}$ by S where $g^{MFN} \in \mathcal{T}(\emptyset)$. Letting $A^i > A^j > A^k$, the only simultaneous farsighted improving path is $g^{MFN} \rightarrow_c (\emptyset)$ where $c = i$ when $i \in S$ and $c = j$ otherwise. This deters S and implies $(\emptyset) \subseteq \mathcal{C}^{*MFN}(\emptyset)$. Since there is no simultaneous farsighted improving path from $T(\emptyset) = (\emptyset)$ and $T(\emptyset) = (\emptyset)$ is obtainable from any $T(\emptyset) = g^{MFN}$ and improving for either i or j, then $g^{MFN} \notin \mathcal{C}^{*MFN}(\emptyset)$. Thus, $(\emptyset) = \mathcal{C}^{*MFN}(\emptyset)$ and $(\emptyset)$ is the unique FDNE.■

Proof of Proposition 5.3. Given Proposition 5.1, two things need to be shown. First, under the multilateralism game, the interval $[\underline{\alpha}_{ms}, \bar{\alpha}_{ms}]$ over which free trade is initially attained must be i) nonempty and ii) shrink as α_{ls} increases from $3 - 5\tau - \varepsilon$ to $3 - 5\tau$, for some small $\varepsilon > 0$. Both follow directly from the condition $\alpha_{ls} < \alpha_{ms} + 1 - 3\tau$ of Proposition 5.2. Second, holding β (and τ) fixed, the interval $[\underline{\alpha}_{ms}, \bar{\alpha}_{ms}]$ over which free trade can be attained under the bilateralism game must become larger. By Proposition 5.1, Proposition 4.5 applies. Moreover, by Proposition 4.5 and the nontriviality condition, free trade can be attained under the bilateralism game for any feasible α_{ms} when $\alpha_{ls} = 3 - 5\tau$ while the same is not true when $\alpha_{ls} = 3 - 5\tau - \varepsilon$. ■

Proof of Proposition 5.4. This follows from Remark B.2 for m and s. For l, it follows from $\pi_l^H > \pi_l^{I,m} > \max\left(\pi_l^{FT}, \pi_l^N\right)$.■

Proof of Proposition 5.5. When BAs are weak stumbling blocs or strong stumbling blocs the relevant unique FDNE of the bilateralism game are (sl), (ml), (ml, sm) and $(\emptyset)$. With respect to (sl) and (ml) the claim of the proposition follows Lemma B.2 and $\pi_k^{FT} > \pi_k^N$. With respect to (ml, sm) it follows from $\pi_m^H > \pi_m^{I,l} > \pi_m^{FT} > \pi_m^N$, $\pi_l^{I,m} > \pi_l^{K,s} > \pi_l^{FT}$, $\pi_l^{FT} \lessgtr \pi_l^{K,s}$ and $\pi_s^{FT} > \max\left(\pi_s^N, \pi_s^{K,l}\right) > \pi_s^O$. With respect to $(\emptyset)$ it follows from Proposition 5.2. When BAs are weak building blocs the relevant unique FDNE is (ml, sl, sm). The claim of the proposition then follows from $\pi_l^H > \pi_l^{I,m} > \pi_l^{FT}$, $\pi_m^{I,l} > \pi_m^{FT} > \pi_m^{K,s}$ and $\pi_s^{FT} > \pi_s^{K,m} > \pi_s^O$.■

C Figures

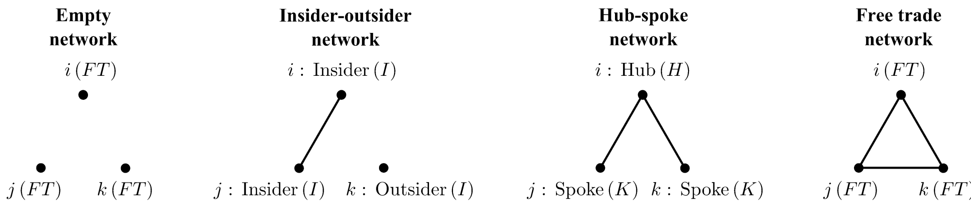

Figure 1: Network positions

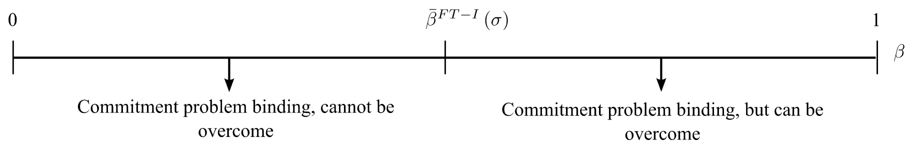

Figure 2: Commitment problem — symmetry

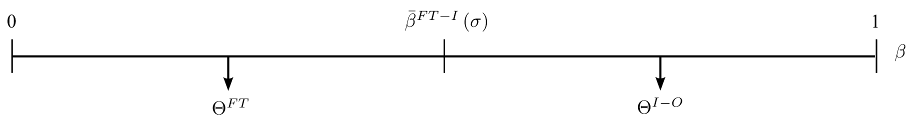

Figure 3: FDNE — symmetry

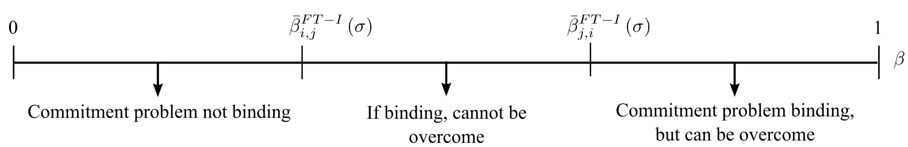

Figure 4: Commitment problem — similar countries

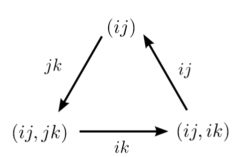

Figure 5: Multiple FDNE

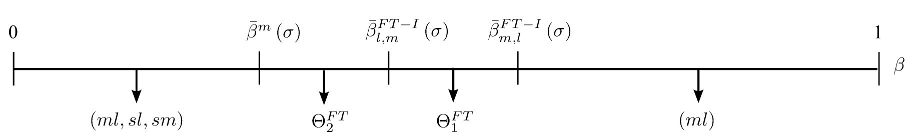

Figure 6: FDNE — similar countries

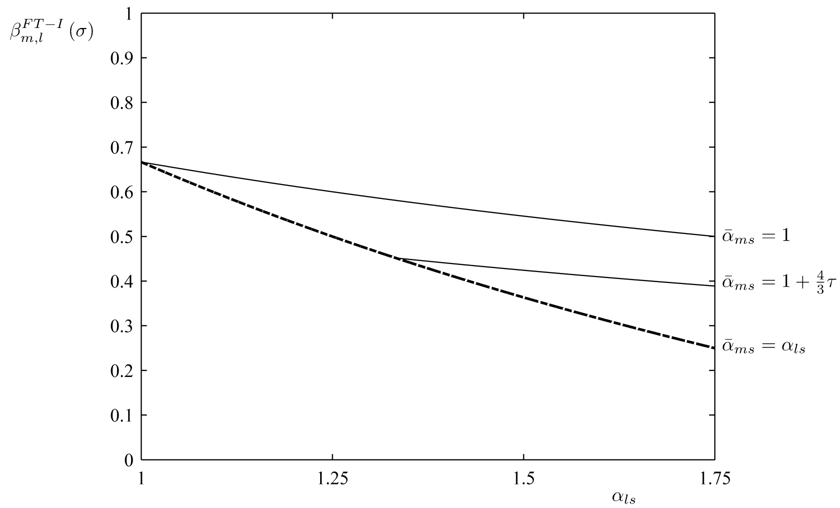

Figure 7: Effect of asymmetry — similar countries

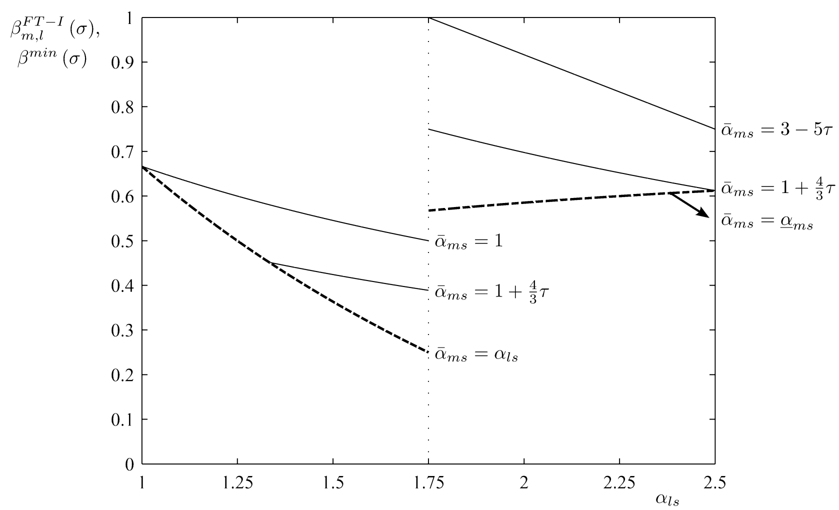

Figure 8: Commitment problem breakdown and effect of asymmetry

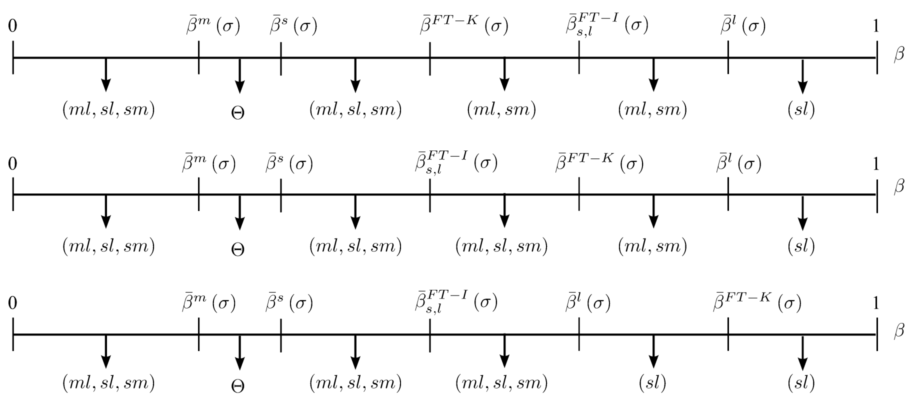

Figure 9: FDNE — commitment problem broken down

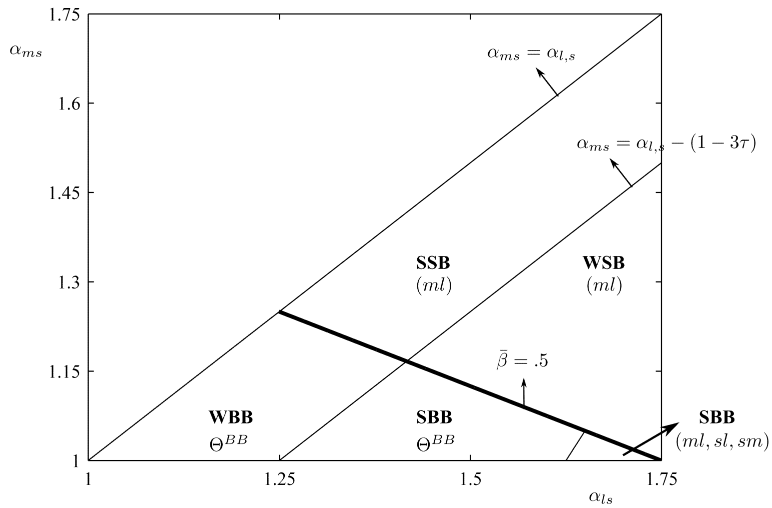

Figure 10: Bilateralism vs multilateralism FDNE — similar countries

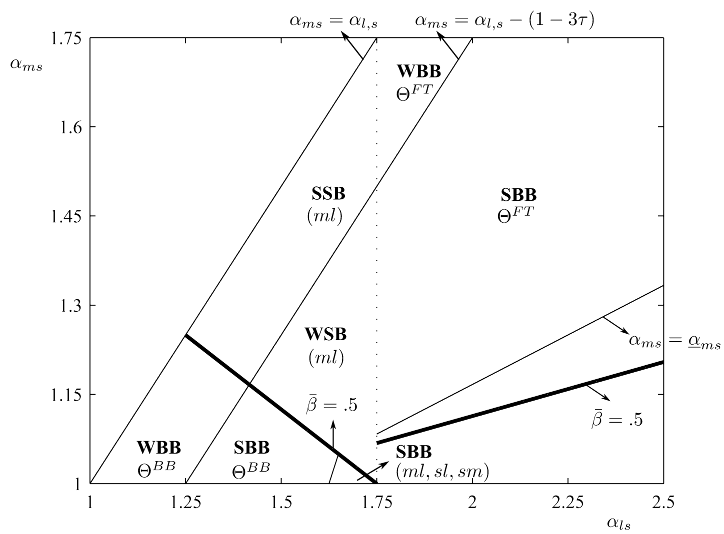

Figure 11: Bilateralism vs multilateralism FDNE — commitment problem breakdown

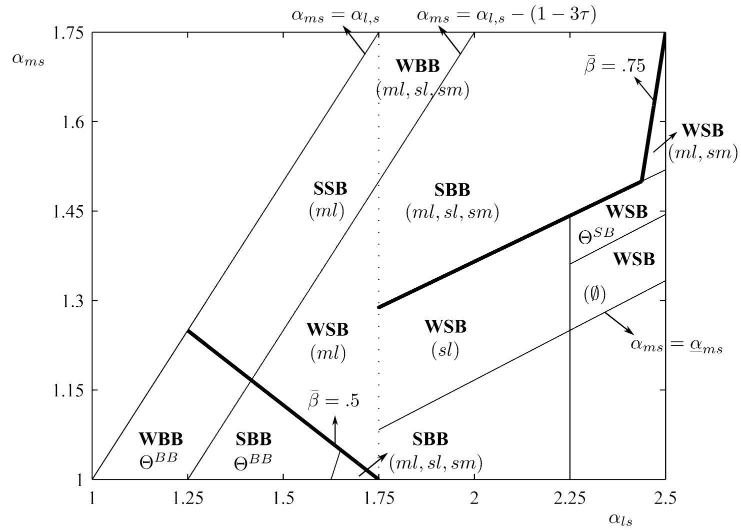

Figure 12: Bilateralism vs multilateralism FDNE — commitment problem breakdown

References

[1] Aghion, P., Antras, P., Helpman, E., 2007. Negotiating free trade. Journal of International Economics 73, 1–30.

[2] Bagwell, K., Staiger, R.W., 1997. Multilateral tariff cooperation during the formation of free trade areas. International Economic Review 38(2), 291–319.

[3] Baldwin, R.E., 1996. A domino theory of regionalism, in: R.E., Haaparanta, P.J., Kiander, J. (Eds.), Expanding the Membership of the EU. Cambridge University Press, Cambridge pp. 25–53.

[4] Baldwin, R.E., 2008. Big–think regionalism: a critical survey. National Bureau of Economic Research Working Paper vol. 14056.

[5] Bernheim, B.D., Peleg, B., Whinston, M.D., 1987. Coalition–proof Nash equilibria I. concepts. Journal of Economic Theory 42, 1–12.

[6] Bhagwati, J.N., 1991. The world trading system at risk. Princeton University Press, Princeton.

[7] Bhagwati, J.N., 1993. Regionalism and multilateralism: an overview, in: Panagariya, A., De Melo J. (Eds.), New Dimensions in Regional Integration. World Bank, Washington D.C., pp. 22–51.

[8] Bond, E.W., Syropoulos, C., Winters, L.A., 2001. Deepening of regional integration and multilateral trade agreements. Journal of International Economics 53, 335–362.

[9] Chwe, M.S., 1994. Farsighted coalitional stability. Journal of Economic Theory 63, 299–325.

[10] Dutta, B., Ghosal, S., Ray, D., 2005. Farsighted network formation. Journal of Economic Theory 122, 143–164.

[11] Ethier, W.J., 2004. Political externalities, nondiscrimination, and a multilateral world. Review of International Economics 12(3), 303–320.

[12] Freund, C., 2000. Multilateralism and the endogenous formation of preferential trade agreements. Journal of International Economics 52, 359–376.

[13] Furusawa, T., Konishi, H., 2007. Free trade networks. Journal of International Economics 72, 310–335.

[14] Goyal, S., Joshi, S., 2006. Bilateralism and free trade. International Economic Review 47(3), 749–778.

[15] Jackson, M.O., 2005. A survey of network formation models: stability and efficiency, in: Demange, G., Wooders, M. (Eds.), Group formation in economics: networks, clubs and coalitions. Cambridge University Press, Cambridge.

[16] Jackson, M.O., 2008. Social and economic networks. Princeton University Press, Princeton.

[17] Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71, 44–74.

[18] Krishna, P., 1998. Regionalism and multilateralism: a political economy approach. Quarterly Journal of Economics 113, 227–251.

[19] Levy, P.I., 1997. A political–economic analysis of free–trade agreements. American Economic Review 87(4), 506–519.

[20] Melatos, M., Woodland, A., 2007. Endogenous trade bloc formation in an asymmetric world. European Economic Review 51, 901–924.

[21] Mukunoki, H., Tachi, K., 2006. Multilateralism and hub and spoke bilateralism. Review of International Economics 14(4), 658–674.

[22] Ornelas, E., 2005. Endogenous free trade agreements and the multilateral trading system. Journal of International Economics 67(2), 471–497.

[23] Ornelas, E., 2008. Feasible multilateralism and the effects of regionalism. Journal of International Economics 74(1), 202–224.

[24] Page Jr., F.H., Wooders, M., Kamat, S., 2005. Networks and farsighted stability. Journal of Economic Theory 120(2), 257–269.

[25] Riezman, R., 1999. Can bilateral trade agreements help to induce free trade? The Canadian Journal of Economics 32(3), 751–766.

[26] Saggi, K., 2006. Preferential trade agreements and multilateral tariff cooperation. International Economic Review 47(1), 29–57.

[27] Saggi, K., Yildiz, H.M., 2010. Bilateralism, multilateralism and the quest for global free trade. Journal of International Economics 81, 26–37.

[28] Saggi, K., Yildiz H.M., 2011. Bilateral trade agreements and the feasibility of multilateral free trade. Review of International Economics. Forthcoming.

[29] Saggi, K., Woodland, A., Yildiz, H.M.. 2010. On the relationship between preferential and multilateral Trade liberalization: the case of customs unions. Mimeo.

[30] Summers, L.H., 1991. Regionalism and the world trading system, in: Policy Implications of Trade and Currency Zones. Federal Reserve Bank, Kansas City.