Why don't more countries form Customs Unions instead of Free Trade Agreements? The role of flexibility

James Lake*
Southern Methodist University

February 14, 2014

Abstract

The vast majority of Preferential Trade Agreements (PTAs) formed are actually Free Trade Agreements (FTAs) rather than Customs Unions (CUs). This largely undocumented prevalence of FTAs relative to CUs is surprising given the traditional view of the literature is that PTA members should prefer CU formation over FTA formation because CU members gain a coordination benefit via setting a common external tariff. I suggest a novel explanation for this prevalence: FTAs possess a dynamic flexibility benefit because individual FTA members have the flexibility to form their own subsequent agreements whereas CU members must jointly engage in any future PTA formation. When the trade off between the CU coordination benefit and FTA flexibility benefit drives the equilibrium, I show that a necessary and sufficient condition for multiple FTAs in equilibrium is that the FTA flexibility benefit dominates the CU coordination benefit. I also use the model to investigate the long standing building bloc–stumbling bloc issue.

JEL codes: C71, F12, F13

Keywords: Preferential Trade Agreement, Free Trade Agreement, Customs Union, coordination, flexibility, free riding, networks, farsighted

*email: jlake@smu.edu. I would like to thank Hiroshi Mukunoki, Murat Yildiz, Alan Woodland, Jim Markusen, Richard Chisik, Christis Tombazos, Taiji Furusawa, Martin Richardson, Maggie Chen and Argya Ghosh for useful comments and discussion as well seminar and conference participants at Monash University, University of New South Wales, Ryerson University, 2013 Australasian Trade Workshop and Spring 2013 Midwest Trade Meetings.

1 Introduction

Since the early 1990s, the world has seen unprecedented growth in the formation of Preferential Trade Agreements (PTAs). According to the WTO (2011, Figure B.1), the number of PTAs increased from around 50 in the late 1980s to nearly 300 by 2010. This trend has spawned numerous strands of literature spanning empirical contributions, e.g. what characteristics determine PTA partners (e.g. Baier and Bergstrand (2004) and Chen and Joshi (2010)), and theoretical contributions, e.g. whether PTAs are "building blocs" or "stumbling blocs" en route to global free trade (Bhagwati (1991)). However, strikingly, Free Trade Agreements (FTAs) outnumber Customs Unions (CUs) by a ratio of 9:1 with the WTO (2011, p.6) listing this phenomenon as one their five stylized facts regarding PTA formation.[1,2] However, as recently argued by Melatos and Woodland (2007, p.904) and Facchini et al. (2012, p.136), the lack of literature explaining this fact is surprising because the existing literature largely suggests CUs are the optimal form of PTA for members.

Unsurprisingly, the standard reason for the attractiveness of a CU relative to an FTA rests on a coordination benefit whereby CU members coordinate their external tariffs.[3] However, the requirement that CU members set a common external tariff implies that individual CU members do not have the flexibility to form their own subsequent PTAs.[4] By formulating a three country dynamic farsighted model where trade agreements form over time, I highlight a novel dynamic benefit of FTAs which helps explain the prevalence of FTAs relative to CUs: FTAs are attractive relative to CUs because they allow individual FTA members to form their own subsequent agreements. Thus, FTAs offer members a *flexibility benefit* while CUs offer members a *coordination benefit*. Indeed, the notion of an FTA flexibility benefit has permeated the mainstream media. Some have argued that the common external tariff of the MERCOSUR CU has prevented Uruguay from forming an FTA with the US. Similar arguments have been made in that the UK and Turkey should have FTAs rather than CUs with the EU to exploit the FTA flexibility benefit.[5]

[1]FTAs differ from CUs because FTA members individually set their tariffs on non–members while CU members set common tariffs on non–members.

[2]`http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx`. Specifically, CUs comprise only 14 of the 180 PTAs that were in force and notified under GATT Article XXIV by 2001 (FTAs make up the remainder) and only 5 of the 169 PTAs formed since 2000.

[3]This coordination benefit could arise from CU members pooling their market power and gaining larger terms of trade gains from non–members (e.g. Bagwell and Staiger (1997a)). It could also arise from, as in this paper, CU members internalizing the negative externality caused by the well known tariff complementarity effect (see Section 2).

[4]If an individual CU member forms a PTA with a nonmember then these two countries eliminate tariffs between themselves. But then the other CU members still have nonzero tariffs with the nonmember which violates the common external tariff.

[5]For the Uruguay case, see http://en.mercopress.com/2011/03/11/how-argentina-

The primary goal of the paper is to analyze how the equilibrium choice between FTAs and CUs depends on the tradeoff between the FTA flexibility benefit and the CU coordination benefit and whether this can help explain the observed prevalence of FTAs relative to CUs. Interestingly, the tradeoff between the FTA flexibility and CU coordination benefits plays a meaningful role only when there is a moderate degree of market size asymmetry (the only source of asymmetry in the model). To this end, it is useful to note that Saggi and Yildiz (2010, p.27) argue "... while Krugman (1991) and Grossman and Helpman (1995) noted that asymmetries across countries can play a crucial role in determining incentives for bilateral and multilateral trade liberalization, existing literature has tended to pay little attention to this issue". For sufficiently small asymmetry, a direct move to global free trade emerges because the fear of being discriminated against as a PTA outsider deters any deviation. For sufficiently large asymmetry, the CU coordination benefit always dominates the FTA flexibility benefit for the two largest countries (the value of coordinating tariffs to protect their large markets is greater than gaining additional access to a small market) and they opt for CU formation.

However, the tradeoff between the FTA flexibility and CU coordination benefits drives PTA formation under a moderate degree of asymmetry. A sufficient condition for CU formation in equilibrium is that the CU coordination benefit dominates the FTA flexibility benefit. Here, CU formation occurs for the same reason as under a large degree of asymmetry. Conversely, a necessary and sufficient condition for multiple FTAs in equilibrium is that the FTA flexibility benefit dominates the CU coordination benefit. In this case, the two largest countries form the first FTA and the largest country becomes the hub on the path to global free trade. Intuitively, having sole preferential market access to the two spoke countries as the hub on the path to global free trade provides enough incentive to the largest country that it forgoes the benefit of coordinating tariffs on the smallest country as part of a CU with the medium sized country. Thus, the model suggests the FTA flexibility benefit helps explain the observed prevalence of FTAs relative to CUs.

Interestingly, FTA free riding incentives that lie off the equilibrium path help ensure multiple FTAs emerge as the *unique* equilibrium when the FTA flexibility benefit dominates the CU coordination benefit. For this uniqueness, the medium sized country must prefer FTA formation with the largest country over PTA formation with the smallest country.

torpedoed-uruguay-s-fta-with-the-us-according-to-wikileaks. For the UK case, see http://blogs.telegraph.co.uk/news/danielhannan/100186074/the-eu-is-not-a-free-trade-area-but-a-customs-union-until-we-understand-the-difference-the-debate-about-our-membership-is-meaningless/. For the Turkish case, see, for example, http://english.alarabiya.net/en/business/economy/2013/05/26/Turkey-fears-being-left-out-in-the-cold-by-EU-free-trade-deals-.html. The Turkish case is somewhat different in that, as part of its CU with the EU, and perhaps in anticipation of EU membership, Turkey agreed to extend any external tariff concessions to future FTA partners of the EU.

Naturally, the larger market access makes the largest country an attractive partner. But, by not forming a PTA with the smallest country, the medium country forgoes a CU coordination benefit as well as the possibility of becoming the hub. When the largest country free rides on the off equilibrium path FTA between the smallest and medium sized countries, the medium country's possibility of becoming the hub vanishes and it prefers an FTA with the largest country over any PTA with the smallest country. This is a noteworthy, and perhaps counter–intuitive, result of the model especially given that FTA formation but not CU formation leads to global free trade.

In addition to providing a novel rationale for the largely unaddressed issue of why so many countries form FTAs rather than CUs, the model provides insights on the more traditional issue of whether PTAs facilitate or hinder the (eventual) attainment of global free trade. I address this "building bloc–stumbling bloc" issue (Bhagwati (1991)) by following the methodological approach emphasized recently by Saggi and Yildiz (2010, 2011) and Saggi et al. (2013) which compares the equilibrium of a game where countries choose between PTAs and MFN agreements and a game where PTAs are not available. While the results largely reflect the existing literature, a number of subtle insights emerge.

First, the model explains the circumstances when PTAs prevent global free trade and when they do not in a setting where the type of PTA (i.e. FTA or CU) is endogenous. Recent results emphasizing CUs can prevent global free trade (e.g. Saggi et al. (2013)) and emphasizing that FTAs do not prevent global free trade (e.g. Saggi and Yildiz (2010)) do so in a setting where countries do not choose the type of PTA. In contrast, the model says that PTAs will prevent global free trade if the CU coordination benefit outweighs the FTA flexibility benefit but not when the FTA flexibility benefit outweighs the CU coordination benefit. In particular, FTAs rather than CUs can emerge in equilibrium and lead to global free trade even when CU members hold a "CU exclusion incentive" meaning they exclude the non member from global free trade. This result emphasizes the importance of endogenizing the choice of PTA type because the CU exclusion incentive drives the result of Saggi et al. (2013) that CUs can prevent global free trade. The result also emphasizes the importance of dynamics because if Missios et al. (2013) allowed countries to choose between CUs and FTAs in their static model (they only compare the equilibrium outcomes of a CU formation game and FTA formation game), FTAs would never emerge.

Second, the model provides the novel insight that FTAs play an important role in *limiting* the destructive nature of PTAs. Specifically, there are cases where i) if CUs were the only type of PTA then the unique equilibrium is a CU between the two largest countries that would not expand to global free trade, yet ii) FTAs emerge on the path to global free trade in the unique equilibrium when the choice between CUs and FTAs is endogenous. The role played by FTAs in limiting the destructive nature of PTAs arises because of the FTA flexibility benefit. Moreover, given the above discussion regarding free riding off the equilibrium path, this role played by FTAs strengthens when there are greater FTA free riding incentives.

Third, the fact that PTAs can prevent and are *never* necessary for global free trade is an unusually strong result because it says that, *for any* degree of market size asymmetry, the prospects for global free trade are highest when countries cannot form PTAs. This result highlights the importance of endogenously determining the equilibrium type of PTA. As discussed later in the introduction, this endogenous determination explains what drives the difference between the negative role of PTAs here relative to other recent papers emphasizing a positive role for a particular type of PTA.

This paper fits into a small literature analyzing the endogenous choice between CUs and FTAs (all such papers are three country models). Papers adopting static approaches include Riezman (1999), Melatos and Woodland (2007) and Facchini et al. (2012). While Riezman (1999) finds CU formation emerges when there are two large countries and one small country (because, like here, such countries have a "CU exclusion incentive"), FTAs never emerge in equilibrium. Similarly, Melatos and Woodland (2007) find FTAs never emerge in a unique equilibrium even though greater preference or endowment asymmetries between countries increase the relative attractiveness of FTAs compared to CUs. Conversely, Facchini et al. (2012) find that PTAs rather than MFN agreements emerge in equilibrium when income inequality is low with FTAs emerging when cross country production structures differ sufficiently. Because of their static nature, none of these papers address the flexibility versus coordination issue at the heart of this paper and only Facchini et al. (2012) addresses the prevalence of FTAs.

Unlike these static models, Seidmann (2009) and Melatos and Dunn (2013) develop dynamic models with an endogenous choice between FTAs and CUs. Melatos and Dunn (2013) argue FTA formation may arise if countries anticipate this allows the possibility to form future FTAs with countries who currently do not trade with the rest of the world. While the spirit of their argument applies to recent new WTO members such as China (2001) and Russia (2012), these countries did engage in global trade prior to WTO membership and Russia even formed PTAs notified to the WTO under GATT Article XXIV. In contrast to Melatos and Dunn (2013), I assume all countries participate in global trade in all periods.

In a three country dynamic bargaining model with transfers, Seidmann (2009) shows PTAs can be valuable because of a "strategic positioning" motive. By affecting the outside option of the PTA outsider, PTA members can affect the share of the global free trade pie obtained by themselves. Because exploiting the strategic positioning motive requires direct expansion of the bilateral PTA to global free trade, CUs are more attractive than FTAs in that CU expansion must immediately result in global free trade whereas FTA expansion can produce overlapping FTAs. Thus, while the flexibility of FTAs is a benefit in this paper it mitigates the strategic positioning motive for PTA formation in Seidmann (2009).[6] In turn, the strategic positioning motive cannot explain the prevalence of FTAs.

A key modeling difference between this paper and Seidmann (2009) that drives the different role of FTAs is the absence of transfers. Even though global free trade maximizes world welfare (here and in Seidmann (2009)), global free trade may not result here because transfers are assumed away (as in, e.g., Saggi et al. (2013)). Bagwell and Staiger (2010, p.50) argue that reality is "... positioned somewhere in between the extremes of negotiations over tariffs only and negotiations over tariffs and [transfers]...".[7] In the absence of transfers, countries payoffs come from the discounted value of one period payoffs along the equilibrium path. Thus, the CU coordination benefit derives from the additional one period payoff of CU over FTA formation while FTA formation is valuable because it allows a country to then become the hub and have sole preferential access to each spoke market.

While this paper finds that PTAs are never necessary for global free trade, Ornelas (2007), Saggi and Yildiz (2010) and Saggi et al. (2013) emphasize PTAs can be necessary for global free trade because the embodied discrimination can induce otherwise unwilling countries to participate in global free trade. Using a nearly identical trade model to this paper, Ornelas (2007) shows the discrimination faced by a large country when two smaller countries form a CU can induce the large country to participate in global free trade. However, importantly, this CU between the smaller countries is the only CU considered by Ornelas (2007). In contrast, this paper shows that whenever the largest country wants to block global free trade, the equilibrium CU is actually between the two larger countries and this CU does not expand to global free trade. Saggi and Yildiz (2010) and Saggi et al. (2013) emphasize that a country may block global free trade so it can free ride on the MFN tariff liberalization of others. This possibility significantly increases the potential that PTAs can play a constructive role. However, in this paper, two country MFN liberalization never occurs in equilibrium.

The paper proceeds as follows. Section 2 presents the underlying trade model. Section 3 presents the dynamic game and equilibrium concept. Section 4 characterizes the equilibrium

[6]Indeed, FTAs only emerge in equilibrium in Seidmann (2009) if tariffs are exogenous. But, in this paper, a key trade off between CUs and FTAs arises precisely because CU members endogenously determine a common external tariff

[7]Other papers assuming transfers between countries are Aghion et al. (2007); Ornelas (2008) and Bagwell and Staiger (2010) while other papers assuming away transfers are Riezman (1999); Furusawa and Konishi (2007); Melatos and Woodland (2007); Saggi and Yildiz (2010) and Facchini et al. (2012). Bagwell and Staiger (2010, p.50) additionally state that "While it is not standard for GATT/WTO trade negotiations to involve explicit transfers as part of the agreement, these negotiations do often involve more than just tariff reductions." and Furusawa and Konishi (2007, p.329) state "...feasible amounts of transfer are usually limited in practice.".

in the special case where market size is symmetric across countries while Section 5 characterizes the equilibrium under asymmetric market size. Section 6 applies the results of Sections 4 and 5 to the building bloc–stumbling bloc issue and Section 7 concludes. Appendix B collects the proofs.

2 Underlying trade model

The underlying trade model is a standard oligopolistic intra industry trade model where governments' objective function is national welfare. However, those familiar with the competing exporters model of Bagwell and Staiger (1997b) will notice that the two models yield the same one period payoff rankings across the various network structures under small enough degrees of asymmetry (e.g. see Saggi and Yildiz (2010), Saggi et al. (2013)). Since these payoff rankings drive the results, the results generalize past the oligopolistic model.

In the model used here, a single firm exists in each of the three countries $N = \{i, j, k\}$. q_{ij} denotes country i's exports to country j and q_{jj} denotes the quantity produced by country j for its domestic market. P_j and $Q_j = \sum_{i \in N} q_{ij}$ denote price and aggregate quantity in country j. The linear inverse demand function in country j is $P_j = \alpha_j - Q_j$ where α_j is the measure of country j's market size. Taking tariffs as given and assuming a common and constant marginal cost (normalized to zero) as well as segmented markets, firm i's maximization problem in country j has the standard form as, e.g., Krishna (1998), Ornelas (2005) and Goyal and Joshi (2006):

$$\max_{q_{ij}} \left[(\alpha_j - Q_j) - \tau_{ji} \right] q_{ij}$$

where τ_{ji} denotes the tariff imposed by country j on country i. Naturally, $\tau_{jj} = 0$ while $\tau_{ji} = 0$ if i and j have a PTA or global free trade prevails. Equilibrium quantity produced by country i and sold in country j's market, given a network of trade agreements g, is

$$q^*_{ij}(g) = \frac{1}{4} \left[\alpha_j + (3 - \eta_j(g)) \bar{\tau}_j(g) - 4\tau_{ji}(g) \right]$$

where $\bar{\tau}_j(g)$ is the MFN tariff faced by countries who do not have a PTA with country j and $\eta_j(g)$ is the number of countries that face a zero tariff in country j (including country j itself). Equilibrium export profits for country i in country j are $\pi_{ij}(g) = \left(q^*_{ij}(g) \right)^2$ and country i's total profits from exporting and domestic production are $\pi_i(g) = \sum_{j \in N} \pi_{ij}(g)$.

Country i's MFN tariff $\bar{\tau}_i(g)$ is its network contingent optimal tariff recognizing that tariffs between PTA members are zero. Except as a CU member, each country retains sovereign discretion over $\bar{\tau}_i(g)$. Thus, each non CU member i solves

$$\max_{\bar{\tau}_i} W_i(\bar{\tau}_i; g) \equiv CS_i(\bar{\tau}_i; g) + \pi_i(\bar{\tau}_i; g) + TR_i(\bar{\tau}_i; g)$$

where $CS_i(\cdot)$ and $TR_i(\cdot)$ are the consumer surplus and tariff revenue in country i. Given a network g where country i is not a CU member, its optimal tariff is:

$$\bar{\tau}_i(g) = \frac{3\alpha_i}{11\eta_i(g) - 1}.$$

Thus, $\bar{\tau}_i(g) = \frac{3\alpha_i}{10}$ when i has no FTA partners while $\bar{\tau}_i(g) = \frac{\alpha_i}{7}$ when i has a single FTA partner.

In contrast, I follow the standard approach of the literature (e.g. Saggi et al. (2013)) where, even with asymmetric countries and the absence of transfers, CU members (say i and j) set a common tariff on nonmembers by maximizing joint welfare $W_i(\bar{\tau}_i, \bar{\tau}_j; g) + W_j(\bar{\tau}_i, \bar{\tau}_j; g)$. Although Syropoulos (2002, 2003), and Melatos and Woodland (2007) show this standard approach may not be innocuous, my results merely rely on the one period CU payoff exceeding the one period FTA payoff because of CU tariff coordination. Thus, I abstract from the complications discussed by these authors. Hence, the optimal common external tariff of a CU between countries i and j is:

$$\bar{\tau}_i(g) = \bar{\tau}_j(g) = \frac{5(\alpha_i + \alpha_j)}{38}.$$

As is well known, "tariff complementarity" is a feature of FTA formation in, among other models, the oligopolistic model.[8] That is, the FTA between i and j induces these countries to reduce their tariffs on country k. While the home government has an incentive to shift domestic market profits from foreign firms to domestic firms by raising tariffs, FTA formation increases domestic market competition. Since the domestic firm's markup falls, this profit shifting motive becomes less attractive and the home government lowers its tariff on the nonmember foreign firm. Importantly, the lower post–FTA tariff on the nonmember not only shifts domestic market profits from the domestic to the nonmember firm but also from the new FTA member firm to the nonmember firm. Thus, tariff complementarity creates a negative "loss sharing" externality between FTA members.[9]

CU formation allows members to internalize the loss sharing externality through setting

[8]It is worth noting that tariff complementarity also arises in simple general equilibrium models of trade agreements such as Bond et al. (2004).

[9]The "loss sharing" terminology comes from Chen and Joshi (2010). Among others, Estevadeordal et al. (2008) provide empirical evidence supporting the tariff complementarity phenomenon.

a common external tariff. This coordination benefit underlies why the bilateral CU optimal tariff exceeds the bilateral FTA optimal tariff: $\frac{5(\alpha_i+\alpha_j)}{38} > \frac{\alpha_i}{7}$.[10] Nevertheless, CU members' ability to exploit the coordination benefit is limited because WTO rules prevent CU members raising external tariffs after CU formation.[11] In this model, the constraint never binds for the larger CU member but binds for the smaller CU member, say j, when $\frac{\alpha_i}{\alpha_j} > \frac{32}{25}$. In this case, the common external tariff of the CU is the smaller member's tariff prior to CU formation which is $\frac{3\alpha_j}{10}$ and, thus, only the larger CU member lowers its external tariff upon CU formation.

PTA formation benefits the nonmember country through lower external tariffs since tariff complementarity holds, at least to some extent, for FTA and CU formation. Nevertheless, the nonmember also faces discrimination in each member market. Thus, PTA formation does not necessarily benefit the nonmember on net. However, FTA formation does benefit the nonmember on net because the tariff complementarity effect dominates. Conversely, CU formation hurts the nonmember on net because coordination of the CU external tariff mitigates the tariff complementarity effect so that the discrimination effect dominates.[12]

The discrimination faced by nonmembers translates into preferential access for members and represents the benefit that members receive from PTA formation. Nevertheless, giving preferential access to a partner country means the domestic firm loses domestic market share to the new member foreign firm and this represents the cost of PTA formation. In general, whether a PTA delivers members a positive net benefit depends on the market size parameters. Section 4 will begin by dealing with this issue under symmetric market size. Bu, before doing so, Section 3 presents the dynamic game and equilibrium concepts.

3 Dynamic game and equilibrium concept

3.1 Overview

The dynamic model has two defining features. First, at most one agreement can form in each period (the game starts with no agreements in place). That is, I view a period as the length of time taken to negotiate a single agreement. This is consistent with reality in that the

[10]Formally, this inequality holds iff $\frac{\alpha_i}{\alpha_j} \in \left[\frac{3}{35}, \frac{35}{3}\right]$. The following analysis respects this constraint.

[11]Like Missios et al. (2013), the CU external tariff constraint plays little role here even if it binds. In contrast, see Mrázová et al. (2012) for an economic environment where the constraint can have implications.

[12]Letting Ø denote the case of no agreements and (jk) denote an FTA between countries j and k then, using the formulas in Appendix A, $W_i(jk) - W_i(\emptyset) \propto .3331\left(\alpha_j^2 + \alpha_k^2\right) > 0$. Moreover, letting $\alpha_j > \alpha_k$ and letting and $\left(jk^{CU}\right)$ denote a CU between j and k, $W_i\left(jk^{CU}\right) - W_i(\emptyset) \propto 1.68\alpha_j^2 + 1.32\alpha_k^2 - 3.6\alpha_j\alpha_k < 0$ for $\frac{\alpha_j}{\alpha_k} < 1.673$ when the WTO external tariff constraint binds and, otherwise, $W_i\left(jk^{CU}\right) - W_i(\emptyset) \propto .7243\left(\alpha_j^2 + \alpha_k^2\right) - 1.9114\alpha_j\alpha_k < 0$ for $\frac{\alpha_j}{\alpha_k} < 2.18$. Throughout the subsequent analysis, $\frac{\alpha_j}{\alpha_k} < 1.673$.

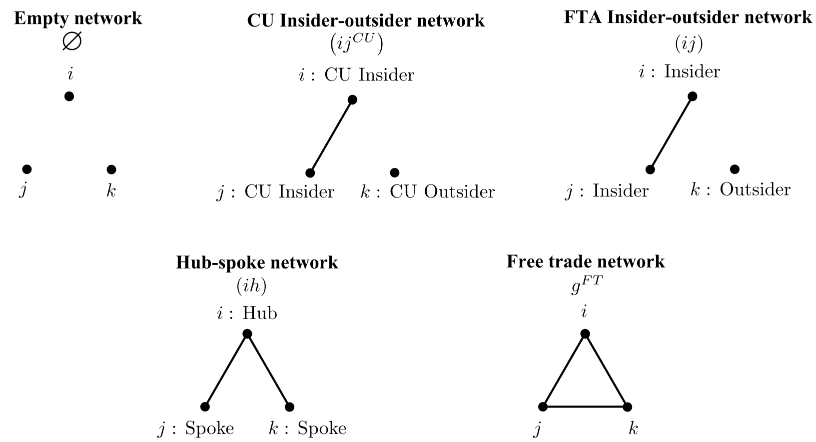

Figure 1: Network positions and notation

length of time taken to conclude a fixed number of agreements is increasing in the number of agreements.[13,14] Since a direct move to global free trade is allowed in any period via a three country MFN agreement, this only rules out formation of two bilateral PTAs in a single period. The second defining feature is that trade agreements formed in previous periods are binding (see last paragraph of this subsection for discussion). Given Markov behavior, this implies the status quo remains forever once no PTA forms in a given period or global free trade is attained. With one agreement per period, the status quo remains after at most three periods.

Figure 1 depicts the possible trade networks and notation. The countries are generically denoted i, j and k and an edge between two countries represents a trade agreement. Two types of situations lead to the free trade network: three individual FTAs or the three country MFN agreement. The three country MFN agreement can occur at the empty network or a CU or FTA insider–outsider network.

As noted by Seidmann (2009), each period can be viewed as a subgame characterized by the trade network that exists at the beginning of the period. Importantly, the payoffs resulting from an outcome in a given subgame are the *continuation* payoffs rather than the one period payoffs. Given the assumption of binding agreements, the model is simply solved using backward induction since global free trade (i.e. the free trade network) is an absorbing state. To determine which agreement forms in a given period, a subgame is modeled as a simultaneous move "announcement game" where each country announces the agreement it wants to form. These announcement games are described in detail in the following subsection.

[13]The length of time between commencement of negotiations and implementation of an agreement is typically many years. For example, NAFTA was implemented in 1994 yet negotiations date back to 1986. Thus, assuming a country could form an unlimited number of agreements in a single period is extreme.

[14]Thus, the discount factor for a period in the model is effectively the one year discount factor raised to a power n where n is the number of years needed to negotiate an agreement.

To solve an announcement game in a given period, I use the solution concept of equilibrium binding agreement (EBA; Ray and Vohra (1997), Diamantoudi (2003)) which is explained in Section 3.3. Importantly, the model does not suffer from the well known difficulty that the equilibrium of sequential move games is often sensitive to the order of negotiations (Ludema (1991); Ray and Vohra (1997); Jackson (2008)). Rather, which countries form an agreement in a given period, and what type of agreement they form, is the outcome of an announcement game that depends endogenously on market size asymmetry.[15]

Before moving on, I make a final observation. Trade agreements formed in previous periods are binding which makes inter–temporal coalitional formation rigid. Conversely, the EBA solution concept assumes a coalition of countries could break up into subcoalitions costlessly while the trade agreement to be formed in the current period is still "under negotiation". I view this dichotomy as a strength rather than a weakness of the model. Many authors (e.g. Ornelas (2008, p.218) and Ornelas and Liu (2012, p.13)) have argued the binding nature of trade agreements is not only pervasive in the trade agreements literature but entirely realistic.[16] Additionally, Lake (2013) describes how Colombia and South Korea began FTA negotiations with Canada before they began negotiations with the US yet both Colombia and South Korea formed FTAs with the US before they did with Canada even though Canada and the US were already part of NAFTA. Thus, I view the contemporaneous flexibility of coalition formation mixed with the inter–temporal rigidity of coalition formation as the appropriate setting for dynamic analyses of trade agreement formation.

3.2 Actions and strategies

The set of countries is generically denoted $N = \{i, j, k\}$. Given the assumption of one agreement per period, each period can be characterized by the network g that exists at the beginning of the period. Given the network at the beginning of a period is g, countries play a simultaneous move "announcement game" to determine which agreement forms in the period. Like Seidmann (2009), I refer to this announcement game as the subgame at network g.

For the subgame at network g, country i's action space $A_i(g)$ represents the set of announcements country i can make. For a coalition $S \subseteq N$, $A_S = \Pi_{i \in S} A_S(g)$ has an analogous interpretation. Given the network notation in Figure 1, Table 1 shows a country's action

[15] This endogenous order contrasts with previous sequential and dynamic models of trade agreements such as Aghion et al. (2007), Mukunoki and Tachi (2006) and Seidmann (2009). The endogenous order also contrasts with the rest of the dynamic network theoretic literature; for example, the model of Dutta et al. (2005).

[16] They argue it is realistic both in terms of real world observation and in terms of being a reduced form shorthand for a more structural modeling approach (see McLaren (2002) for sunk costs as a structural justification and Roberts and Tybout (1997); Eichengreen and Irwin (1998) and Freund and McLaren (1999) for empirical support).

space consists of three types of announcements $a_i(g) \in A_i(g)$. First, the country with whom it wants to form a PTA but has not yet done so (the superscript CU indicates announcement of a CU and absence of a superscript indicates announcement of an FTA). Second, a direct move to the free trade network, denoted FT. Third, no announcement, denoted ϕ. An agreement forms when all members of the proposed agreement announce in favor: the FTA between i and j forms if and only if $a_i(g) = j$ and $a_j(g) = i$ while the three country agreement resulting in global free trade forms if and only if $a_i(g) = a_j(g) = a_k(g) = FT$.

Given the tight link between action profiles and resulting networks, I will often refer to a coalition S deviating from one network to another. Of course, formally speaking, the coalition S deviates from one coalitional action a_S to another a'_S. For example, the coalitional deviation by $S = ij$ from $a_S = (a_i, a_j) = (\phi, \phi)$ to $a'_S = (j, i)$ induces an FTA between i and j and so I refer to this as i and j deviating from $g = \emptyset$ to $g' = (ij)$.[17]

The dynamic nature of dynamic model implies that whether the coalition $S = ij$ is "better off" or "worse off" in the above example under the new action profile a' compared to the initial action profile a depends on the comparison of continuation payoffs associated with g and g' rather than the one period payoffs. Since these continuation payoffs depend on the equilibrium path of trade networks stemming from g and g', it will be useful to let $\langle g \rangle$ and $\langle g' \rangle$ denote these paths with $\langle g' \rangle \succ_S \langle g \rangle$ denoting that each member of S receives a higher continuation payoff under $\langle g' \rangle$ than $\langle g \rangle$. Finally, I let $\langle g \rangle = g$ denote that the equilibrium path of agreements stemming from g actually remains at g forever. For example, $\langle (ij) \rangle = (ij)$ denotes that the FTA (ij) remains forever once it forms.

Network	**Player action space**		
	$A_i(g)$	$A_j(g)$	$A_k(g)$
$\emptyset$	$\{\phi, j, j^{CU}, k, k^{CU}, FT\}$	$\{\phi, i, i^{CU}, k, k^{CU}, FT\}$	$\{\phi, i, i^{CU}, j, j^{CU}, FT\}$
(ij^{CU})	$\{\phi, FT\}$	$\{\phi, FT\}$	$\{\phi, FT\}$
(ij)	$\{\phi, k, FT\}$	$\{\phi, k, FT\}$	$\{\phi, i, j, FT\}$
(ih)	$\{\phi\}$	$\{\phi, k\}$	$\{\phi, j\}$
g^{FT}	$\{\phi\}$	$\{\phi\}$	$\{\phi\}$

Table 1: Action space for each subgame

3.3 Equilibrium concept used to solve each subgame

The solution concept I use to solve the announcement game at each network (i.e. to solve each subgame) is equilibrium binding agreement (EBA; Ray and Vohra (1997) and Diamantoudi

[17] Of course, formally, $S = ij$ is shorthand for $S = \{i, j\}$.

(2003)). Having solved for the EBA in each subgame I will simply refer to the resulting equilibrium path of agreements as *the equilibrium path of agreements.*

The key idea behind an EBA is that a deviating coalition does not take the actions of other players as given but rather anticipates equilibrium reactions of other players. This idea is formalized as follows. Take an initial action profile a. Then, the action profile $a' = (a'_S, a'_{N \backslash S})$ resulting after S deviates to a'_S (to be clear, $a'_{N \backslash S} = a_{N \backslash S}$ is *not* imposed) must satisfy two properties :

1. a' is a Nash equilibrium between S as one player and all other players, $N \setminus S$, as the second player.[18]

2. For any unilateral deviation by some player i from a'_i to a''_i there is a Nash equilibrium $a'' = \left(a''_i, a''_{-i}\right)$ that leaves i no better off than under a'.[19,20]

An action profile a' that satisfies properties 1 and 2 is an EBA between S and $N \setminus S$. A deviation by S from a_S to a'_S is self enforcing if, for any $a' = \left(a'_S, a'_{N \backslash S}\right)$ that is an EBA between S and $N \setminus S$, each member of S is better off under a' relative to a. Intuitively, a deviation by S is self enforcing if S wants to deviate given it anticipates (any) "equilibrium reactions" of the other players.[21] An action profile a is an EBA if it is Pareto optimal (i.e. it is a Nash equilibrium for N) and there is no self enforcing deviation by a coalition $S \subset N$. However, there may exist a self enforcing deviation from every action profile. In this case, an action profile is an EBA if it is an EBA between some S and $N \setminus S$ or, equivalently, it satisfies properties 1 and 2 for some S. If there is no EBA between any S and $N \setminus S$ then the EBAs are the Nash equilibria. Although, formally, an EBA is an action profile, I will refer to the agreement induced by such an EBA action profile as an EBA. Also, I will refer to an agreement induced by a Nash equilibrium as a Nash agreement.

The idea that a deviating coalition anticipates the equilibrium reactions of other players rather than taking the actions of other players as fixed is the key difference between an EBA

[18]For clarity, suppose S is a two player coalition. Then, $a' = (a'_S, a'_{N \backslash S})$ is a Nash equilibrium between S and $N \setminus S$ iff i) there is no $a'' = \left(a''_S, a'_{N \backslash S}\right)$ that makes each member of S better off relative to a' and ii) there is no $a'' = (a'_S, a''_i)$ that makes $N \setminus S \equiv i$ better off relative to a'.

[19]Given property 1 and the three player context, only unilateral deviations by members of the two player coalition (which could be either S or $N \setminus S$) need consideration in property 2.

[20]To be clear, $a''_{-i} = a'_{-i}$ is *not* imposed.

[21]For a given a'_S, there can be many action profiles $a' = (a'_S, a'_{N \backslash S})$ satisfying properties 1 and 2. The definition says the deviating coalition S undertakes their deviation only if they prefer *any* such a' over the initial action profile a. This was proposed by Diamantoudi (2003) and differs from Ray and Vohra (1997) who require S merely prefer *some* such a' to a. Essentially, Diamantoudi (2003) assumes deviating coalitions anticipate "pessimistically" while Ray and Vohra (1997) assume they anticipate "optimistically".

and the well known concept of coalition proof Nash equilibrium (CPNE).[22] Nevertheless, despite the inherent logical appeal, an EBA is more complicated than a CPNE. Thus, why not solve a CPNE rather an EBA in every subgame? Indeed Bernheim et al. (1987) define this as a Perfectly CPNE. The fundamental problem is that CPNE non–existence arises in simple settings such as Condorcet paradox type situations and these situations arise in the model. This questions the fundamental validity of CPNE to explain strategic formation of trade agreements in dynamic contexts and naturally leads to similar but stronger concepts such as EBA.

For those reading the proofs in Appendix B, I will begin Appendix B by outlining some notation that helps facilitate exposition of the proofs and will *only* be used in the proofs.

3.4 Stylized examples

The following two (highly stylized) examples illustrate the process of deriving an EBA in a particular subgame and the process of deriving the equilibrium path of agreements using backward induction.

Example 1. The example derives the EBA for the subgame at the CU insider–outsider network $g_0 = \left(ml^{CU}\right)$. That is, the network at the beginning of the period is g_0. The two possible agreements (i.e. outcomes) for the subgame are no agreement, $\emptyset$, and a direct move to global free trade, g^{FT}, which yield the respective networks $g_1 = \left(ml^{CU}\right)$ or $g_1 = g^{FT}$.[23] Since either network is an absorbing state, player i's continuation payoff for network g_1 normalized by $1-\beta$ (where β is the discount factor) is just the one period payoff $W_i\left(g_1\right)$.

The example is highly simplified because remaining CU insiders is Pareto dominant for m and l. This Pareto dominance implies m and l have a self enforcing deviation from g^{FT} to $\emptyset$. $\emptyset$ is clearly Nash between $S = ml$ and s (property 1) and the Pareto dominance implies neither m nor l will subsequently deviate (property 2). Thus, $\emptyset$ is an EBA between $S = ml$ and s. Given the Pareto dominance, only s wants to deviate from $\emptyset$. However, s has no such self enforcing deviation because g^{FT} is not an EBA between $S = ml$ and s (violation of property 2) since m or l would subsequently deviate to $\emptyset$ given it is the unique Nash agreement. Indeed, g^{FT} is not even Nash between $S = ml$ and s (violation of property 1).

[22] Another important similarity between EBA and CPNE is that both concepts view "stability" of an action profile not with respect to all possible deviations but only those that are themselves equilibria. This idea emerges above because an action profile satisfying properties 1 and 2 is an EBA between S and $N \setminus S$. The idea emerges in CPNE where, in the three player context (for example), consideration is given to whether a bilateral deviation would then induce a unilateral deviation by one of the initial deviators.

[23] Note, $\emptyset$ denotes the empty network meaning no agreements exist. Conversely, $\emptyset$ denotes the outcome in a subgame is that no agreement forms.

Thus, given $\emptyset$ is Pareto optimal, it is the unique EBA of the subgame meaning that m and l remain CU insiders.■

$\emptyset$	$\left(ml^{CU}\right)$	$\left(sl^{CU}\right)$	$\left(sm^{CU}\right)$	(ml)	(sl)	(sm)	(lh)	(mh)	(sh)	FT
$(0,0,0)$	$(0,8,8)$	$(7,0,7)$	$(6,6,0)$	$(4,4,4)$	$(3,3,3)$	$(2,2,2)$	$(3,3,9)$	$(1,8,1)$	$(7,1,1)$	$(5,5,5)$

Table 2: Network dependent one period payoffs $(W_s(g), W_m(g), W_l(g))$.

Example 1 illustrates that countries remain CU insiders when doing so is Pareto dominant for each member. Lemma 1 generalizes this intuition: Pareto dominant agreements of a subgame that can be sustained by coalitions are EBAs of the subgame. Following custom, $g+\ell$ denotes that the agreement ℓ is added to network g.

Lemma 1. *Suppose the network at the beginning of period t is g_t and fix $a(g)$ for all $g \neq g_t$. An agreement ℓ_t is an EBA of the subgame in period t if i) $\langle g_t+\ell_t\rangle = \langle \ell_t\rangle = \left\langle g^{FT}\right\rangle$ is Pareto dominant for N, ii) $\langle g_t+\ell_t\rangle = \langle g_t\rangle$ is Pareto dominant for i and $A_i(g)=\{\phi, FT\}$, or iii) $\langle g_t+\ell_t\rangle = \langle g_t+ij\rangle$ or $\langle g_t+\ell_t\rangle = \langle g_t\rangle$ is Pareto dominant for i and j. If Pareto dominance is strict, then ℓ_t is the unique EBA of the subgame in period t.*

Example 2 illustrates the process of deriving the equilibrium path of agreements by backward induction.

Example 2. Since the free trade network is an absorbing state, consider the subgame at the hub–spoke network $g_0=(lh)$. That is, the network at the beginning of the period is $g_0=(lh)$. The two possible EBAs in the subgame are no agreement, $\emptyset$, or the spoke–spoke FTA (sm). Respectively, these agreements yield an end of period network, g_1, of (lh) or g^{FT}. In either case, g_1 is an absorbing state. Thus, using Table 2, s and m have the same continuation payoffs across either possible g_1: $\frac{3}{1-\beta}$ and $\frac{5}{1-\beta}$ respectively. Hence, $g_1=g^{FT}$ is strictly Pareto dominant for s and m and Lemma 1 implies the unique EBA is the spoke–spoke FTA (sm). Given Table 2, the same logic applies for any hub–spoke network meaning any hub–spoke network expands to the free trade network.

Now consider the subgame at an FTA insider–outsider network $g_0=(ij)$. The four possible EBAs in the subgame are $\emptyset$, (ik), (jk) and g^{FT}. The respective end of period networks, g_1, are (ij), (ih), (jh) and g^{FT}. For illustration, assume the larger insider becomes the hub in an EBA.

Now consider the subgame at a CU insider–outsider network $g_0=\left(ij^{CU}\right)$. For $g_0=\left(ml^{CU}\right)$, Example 1 illustrates that the EBA is $\emptyset$ meaning that m and l remain CU insiders. Using Table 2, similar logic applies for any subgame at a CU insider–outsider network.

Now consider the subgame at the empty network $g_0 = \emptyset$ and let $\beta = \frac{1}{2}$. The eight possible EBAs for the subgame, and thus the end of period network g_1, are $\emptyset$, g^{FT}, (ml), (sl), (sm), $\left(ml^{CU}\right)$, $\left(sl^{CU}\right)$, $\left(sm^{CU}\right)$. Using Table 2, notice that forming a CU is Pareto dominant for m and l unless the normalized continuation payoff associated with forming an FTA and becoming the hub on the path to global free trade exceeds 8. However, l's normalized continuation payoff from forming an FTA with m is $\frac{11}{2}$ (and with s is 5) while m's normalized continuation payoff from forming an FTA with s is 4.[24] Thus, Lemma 1 implies the unique EBA is m and l become CU insiders. Since m and l become CU insiders in period 1 and remain so forever, the unique equilibrium path of agreements is merely $\left\langle \left(ml^{CU}\right) \right\rangle = \left(ml^{CU}\right)$.■

4 Symmetric countries

Although, in general, a member's net benefit from PTA formation depends on the market size parameters, PTA formation yields a positive net benefit when countries are symmetric with one exception. Lemma 2 summarizes.

Lemma 2. *Assume countries are symmetric. Then, i) being a spoke yields lower welfare than being an FTA outsider, i.e.* $W_i(jk) > W_i(kh)$*, but any other bilateral PTA is mutually welfare enhancing, ii) only the hub payoff exceeds the global free trade payoff, i.e.* $W_i(g) > W_i\left(g^{FT}\right)$ *iff* $g = (ih)$*, and iii) welfare is higher for a CU member than an FTA member, i.e.* $W_i\left(ij^{CU}\right) > W_i(ij)$.

Notice that part iii) formalizes the idea that the CU coordination benefit makes CUs more attractive than FTAs from a myopic perspective. Moreover, part ii) creates the possibility of an FTA flexibility benefit since CU members cannot become the hub. Indeed, $W_i(ih) > W_i\left(g^{FT}\right)$ underlies the logic explaining why the FTA flexibility benefit can outweigh the CU coordination benefit.

The first step in using backward induction to solve for the equilibrium path of agreements is solving the EBA for subgames at hub–spoke networks.[25] Given Lemma 2 implies any spoke–spoke FTA is mutually welfare enhancing, Lemma 1 implies this FTA is the EBA for the subgame. Thus, any hub–spoke network expands to global free trade.

---

[24]For example, l's normalized continuation payoff from forming an FTA with m is is $(1-\beta)\left(4+\beta\cdot 9+\frac{\beta^2}{1-\beta}\cdot 5\right) = 4+5\beta-4\beta^2 = \frac{11}{2}$.

[25]Formally speaking, and as mentioned in Sections 3.2 and 3.3, the spoke–spoke FTA is the *EBA outcome* rather than the *EBA* since an EBA is an action profile rather than a trade agreement. This abuse of terminology should not cause confusion.

Rolling back to subgames at FTA insider–outsider networks, a key consideration in determining the EBA is whether the outsider wants to become a spoke by forming an FTA with an insider. Indeed, Lemma 2 says that FTA formation hurts an outsider from a myopic perspective. That is, due to tariff complementarity, $W_i(jk) > W_i(kh)$. However, part ii) of Lemma 2 implies $W_i\left(g^{FT}\right) > W_i(jk)$ and thus an outsider may form an FTA because global free trade eliminates the discrimination faced as an outsider. An outsider, say country i, wants to form an FTA and becomes a spoke if and only if the Free Trade–Outsider (FT–O) condition holds: $W_i(kh) + \frac{\beta}{1-\beta} W_i\left(g^{FT}\right) > \frac{1}{1-\beta} W_i(jk)$. This reduces to

$$\beta > \bar{\beta}_i^{FT-O}(\alpha) \equiv \frac{W_i(jk) - W_i(kh)}{W_i(g^{FT}) - W_i(kh)} \tag{1}$$

where $\alpha \equiv (\alpha_i, \alpha_j, \alpha_k)$. Thus, the outsider wants to become a spoke when $\beta > \bar{\beta}_i^{FT-O}(\alpha)$ because then the absence of discrimination under global free trade outweighs the myopic free riding incentives created by tariff complementarity.

Despite the FT–O condition providing a necessary condition for the emergence of hub–spoke networks in equilibrium, the following lemma shows that, under symmetry, hub–spoke networks never emerge in equilibrium.

Lemma 3. *Assume countries are symmetric and consider the subgame at an FTA insider–outsider network (ij). The unique EBA is a direct move to global free trade g^{FT}.*

Conditional on forming an FTA, $W_i(ih) > W_i\left(g^{FT}\right)$ implies each insider prefers to become the hub rather than move directly to global free trade. However, g^{FT} emerges as the unique EBA because the fear of becoming a spoke or remaining an insider deters an insider from deviating and attempting to become the hub.

More specifically, the following three outcomes are EBAs between $S = jk$ and $N \setminus S = i$: (1) g^{FT}, (2) (jk) if $\beta > \bar{\beta}^{FT-O}(\alpha)$ (i.e. j becomes the hub) and (3) no agreement, denoted $\emptyset$, if $\beta < \bar{\beta}^{FT-O}(\alpha)$ (i.e. i and j remain insiders).[26] This follows because i) all are Nash between $S = jk$ and i and ii) the fear of either (ik) or $\emptyset$ as a Nash agreement deters any deviation by j (the former when $\beta > \bar{\beta}^{FT-O}(\alpha)$ and the latter when $\beta < \bar{\beta}^{FT-O}(\alpha)$) and the fear of $\emptyset$ also deters k's deviation to $a_k(ij) = FT$. Thus, a direct move to global free trade is an EBA because the only profitable deviation is to become the hub but, for example, i's deviation from $a_i(ij) = FT$ to $a_i(ij) = k$ is deterred by (jk) if $\beta > \bar{\beta}^{FT-O}(\alpha)$ and $\emptyset$ if $\beta < \bar{\beta}^{FT-O}(\alpha)$. Moreover, there are no other EBAs because $S = jk$ has a self enforcing deviation to g^{FT} from anything other than g^{FT} or (jk) and, by symmetry, $S = ik$ has a self enforcing deviation from (jk) to g^{FT}.

[26] The country subscript on $\bar{\beta}^{FT-O}(\alpha)$ is dropped because of symmetry.

Indeed, the world not only moves directly from a bilateral FTA to global free trade but also from a bilateral CU to global free trade. Given Lemma 2 implies $W_i\left(g^{FT}\right) > W_i\left(ij^{CU}\right)$, CU members have no "CU exclusion incentive" and expansion of a CU to global free trade is Pareto dominant for all players. Thus, Lemma 1 implies a direct move to global free trade is the EBA in any subgame at a CU insider–outsider network.

Hence, rolling back to the subgame at the empty network, a direct move to global free trade is Pareto dominant for all players. Thus, Proposition 1 follows directly from Lemma 1 and is rather unsurprising given similar results by Saggi and Yildiz (2010, Proposition 1) and Saggi et al. (2013, Proposition 1).

Proposition 1. *Under symmetry, a direct move to global free trade is the unique equilibrium path of agreements.*

5 Asymmetric countries

5.1 "Small" degree of asymmetry

The primary motivation of the paper is to explore the trade–off between the flexibility benefit of FTAs versus the coordination benefit of CUs. Yet, this tradeoff does not play any role in equilibrium when countries are symmetric, in part because any bilateral FTA would immediately expand to global free trade in equilibrium (see Lemma 3). Thus, I now introduce asymmetric market size.

The three countries are the small (s), medium (m) and large (l) countries. Hereafter, asymmetry means the following:

Definition 1 (Asymmetric countries)**.** i) $\alpha_{ls} \equiv \frac{\alpha_l}{\alpha_s} > \alpha_{ms} \equiv \frac{\alpha_m}{\alpha_s} > 1$ and ii) $\alpha_{ls} > \underline{\alpha}_{ls} \equiv 1.01$.

Part ii) places a lower bound on the relative size l. $\alpha_{ls} > \underline{\alpha}_{ls}$ implies $W_s\left(lh\right) > W_s\left(ml\right)$ and $W_m\left(lh\right) > W_m\left(sl\right)$ meaning that, as outsiders, s and m have no free riding incentive since $\bar{\beta}_i^{FT-O}\left(\alpha\right) < 0$ for $i = s, m$. The area of the parameter space analyzed in this section, interpreted as a "small" degree of asymmetry, is summarized by Condition 1. Condition 1 says that CU insiders do not have a CU exclusion incentive since they benefit from including the outsider in expansion to global free trade.

Condition 1 (Small degree of asymmetry)**.** Countries are asymmetric and $W_i\left(g^{FT}\right) > W_i\left(ij^{CU}\right)$ for any i, j.

Condition 1 is depicted in Figure 2 by areas C1(a) and C1(b). Despite asymmetry, Lemma 2 still holds subject to Definition 1. That is:

Lemma 4. *Assume Condition 1 holds. Then,* $W_l(sm) > W_l(kh)$ *for* $k = s, m$ *but any other bilateral PTA is mutually profitable. Moreover, parts ii)-iii) of Lemma 2 hold and a PTA is more attractive with a larger partner.*

Like the symmetric case, the first step in solving the equilibrium path of agreements is simple: spoke–spoke FTA formation is mutually profitable so this FTA is the EBA for subgames at hub–spoke networks and, thus, any hub–spoke network expands to global free trade.

However, unlike the symmetric case, hub–spoke networks can now emerge in equilibrium. This equilibrium emergence of FTAs occurs because of the FTA flexibility benefit. Lemma 5 summarizes.

Lemma 5. *Assume countries are asymmetric and consider the subgame at an FTA insider–outsider network* $g = (ij)$ *where* $\alpha_i > \alpha_j$. *The unique EBA is* g^{FT} *if* $g = (sm)$ *and* $\beta < \bar{\beta}_l^{FT-O}(\alpha)$. *Otherwise, the EBAs are* g^{FT} *and* (ik) *which yields the hub–spoke network* (ih).

When an outsider can credibly refuse to participate in further FTAs, which only happens when l is an outsider and $\beta < \bar{\beta}_l^{FT-O}(\alpha)$, the unique EBA remains a direct move to global free trade. Otherwise, the FTA between the larger insider i and the outsider k is also an EBA which results in the larger insider becoming the hub. Underlying this result is that, unlike the symmetric case, i becomes the hub in the unique Nash agreement by announcing it wants to become the hub and deviating from $a_i(ij) = FT$ to $a_i(ij) = k$. Under symmetry, i feared becoming a spoke in a Nash agreement because of the outsider's indifference regarding the identity of its FTA partner. However, given a PTA is more attractive with a larger partner, the unique Nash agreement is now (ik) when $a_i(ij) = k$.

Like the symmetric case, g^{FT} is an EBA between $S = jk$ and i. Thus, S has a self enforcing deviation to g^{FT} from anything other than g^{FT}. However, (ik) is also an EBA between $S = jk$ and i given the Nash agreement (ik) deters any deviation by j and no agreement, i.e. $\emptyset$, as a Nash agreement deters any deviation by k. Moreover, $\emptyset$ and (jk) are not EBAs between any S and $N \setminus S$ because i or k will deviate given (ik) is the unique Nash agreement conditional on $a_i(ij) = k$ or $a_k(ij) = i$. Thus, i has a self enforcing deviation from g^{FT} and $a_i(ij) = FT$ to $a_i(ij) = k$ and (ik). Hence, there is a self enforcing deviation from any action profile. Thus, the EBAs for the subgame are g^{FT} and (ik) since these are the only EBAs between some S and $N \setminus S$.

This multiplicity of EBAs creates a problem, but not because of the multiplicity itself. Given countries anticipate equilibrium play on and off the equilibrium path, players anticipate the path of agreements conditional on each FTA insider–outsider network. However, multiplicity means countries could anticipate hub–spoke networks or a direct move to global

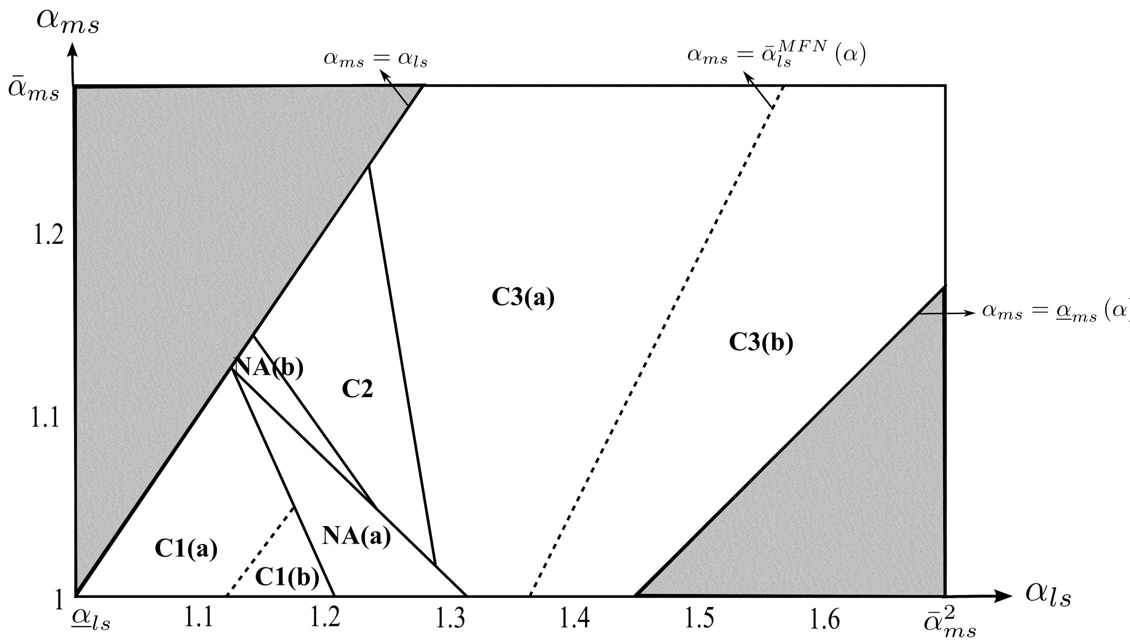

Figure 2: Parameter space

free trade from *each* FTA insider–outsider network. To avoid this complication and to focus on the FTA flexibility benefit, which only exists in the presence of hub–spoke networks, the analysis hereafter ignores g^{FT} as an EBA when there are multiple EBAs in the subgame at the FTA insider–outsider network.[27]

Given the possibility of hub–spoke networks in equilibrium, the FTA flexibility benefit now emerges. This benefit can be seen by ranking, from the large country's perspective, the different equilibrium paths of agreements that emerge depending on the type of PTA that l forms in the first period. In particular, the equilibrium path of agreements conditional on (ml) in the first period is $\langle (ml) \rangle = (ml, sl, sm)$ in contrast to $\left\langle \left(ml^{CU}\right) \right\rangle = \left(ml^{CU}, g^{FT}\right)$ if $\left(ml^{CU}\right)$ forms in the first period. Thus, l prefers to form an FTA rather than a CU with m if and only if

$$W_l\,(ml) + \beta W_l\,(lh) + \frac{\beta^2}{1-\beta} W_l\left(g^{FT}\right) \quad > \quad W_l\left(ml^{CU}\right) + \frac{\beta}{1-\beta} W_l\left(g^{FT}\right). \tag{2}$$

[27]NAFTA provides an *exception* to the rule illustrating the possibility of a bilateral FTA expanding directly into a trilateral FTA rather than the creation of overlapping FTAs. In the NAFTA case, the US and Canada already had a bilateral FTA, known as CUSFTA, which expanded directly into the trilateral FTA known as NAFTA with the inclusion of Mexico.

Rearranging (2) yields:

$$\beta \underbrace{\left(W_l\left(lh\right) - W_l\left(g^{FT}\right)\right)}_{\text{FTA flexibility benefit}} > \underbrace{\left(W_l\left(ml^{CU}\right) - W_l\left(ml\right)\right)}_{\text{CU coordination benefit}}. \tag{3}$$

That is, l prefers FTA formation over CU formation if and only if the FTA flexibility benefit dominates the CU coordination benefit.

The decomposition of (3) into the FTA flexibility benefit and CU coordination benefit is intuitive. $W_l\left(ml^{CU}\right) - W_l\left(ml\right) > 0$ represents the CU coordination benefit which arises because, via external tariff coordination, CU members internalize the negative externality of tariff complementarity. The negative externality is that FTA formation induces m and l to lower their external tariff on s and this shifts domestic market profits in country l from m to s and in country m from l to s. The FTA flexibility benefit arises because FTA formation allows a member to subsequently form its own individual FTA whereas, because of the common external tariff, CU expansion requires the participation of all CU members. $W_l\left(lh\right) - W_l\left(g^{FT}\right) > 0$ represents this flexibility benefit and is valuable because l holds sole preferential access in each spoke market as the hub. Capturing the tradeoff between the immediate CU coordination benefit and the future FTA flexibility benefit, rearranging (3) shows that l prefers an FTA over a CU with m if and only if

$$\beta \quad > \frac{W_l\left(ml^{CU}\right)-W_l(ml)}{W_l(lh)-W_l(g^{FT})} \equiv \bar{\beta}_l^{Flex}\left(\alpha\right) \quad . \tag{4}$$

Rolling back to the empty network, Proposition 2 characterizes the important features of the equilibrium under a small degree of asymmetry. Since a full characterization is rather tedious, the full characterization is presented in Appendix C.

Proposition 2. *Assume Condition 1 holds. Then, a direct move to global free trade, g^{FT}, is an equilibrium path of agreements and is unique if $\alpha_{ls} < \sqrt{.59\left(\alpha_{ms}^2+1\right)}$. Even though this direct move is not always unique, global free trade is eventually attained in any equilibrium path of agreements.*

Like the symmetric case, a direct move to global free trade can be the unique equilibrium path of agreements. Part iii) says a sufficiently small degree of asymmetry is a sufficient condition for uniqueness. Proposition 8 in Appendix C provides the necessary and sufficient conditions and Figure 2 confirms the same interpretation by depicting that, in general, g^{FT} is unique only in area C1(a). The sufficient condition given in part iii) comes from determining whether l prefers a direct move to global free trade over forming an FTA and being the hub on the path to global free trade. While the FTA flexibility benefit is valuable because the hub has sole preferential access in each spoke, $W_i\left(g^{FT}\right) > W_i\left(ij\right)$ implies FTA members do not have an "FTA exclusion incentive" and benefit from expanding their FTA to global free trade. Thus, even the would–be hub prefers a direct move to global free trade unless the FTA flexibility benefit is sufficiently large which, in turn, requires a sufficiently large α_{ls}.

However, unlike the symmetric case, g^{FT} is not always the unique equilibrium path of agreements. Although global free trade is eventually attained in any equilibrium path of agreements, various paths of PTAs can emerge. One path that can emerge is (ml, sl, sm): m and l form the initial FTA and then l is the hub on the path to global free trade. This happens when the following four conditions are satisfied. First, $\langle (ml) \rangle \succ_l \langle g^{FT} \rangle$: the FTA flexibility is large enough that l wants to become the hub on the path to global free trade rather than move directly to global free trade. Second, $\langle (sl) \rangle \succ_l \langle \left(ml^{CU}\right) \rangle$: the FTA flexibility dominates the CU coordination benefit and to the extent that l prefers an FTA with s over a CU with m. Third, $\beta < \bar{\beta}_l^{FT-O}\left(\alpha\right)$: l free rides on the FTA between m and s rather than forming a subsequent FTA with m. Fourth, $\langle (ml) \rangle \succ_m \langle \left(sm^{CU}\right) \rangle$: given m's inability to become the hub after an FTA with s, m prefers to obtain sole preferential access to l's market via an FTA rather than form a PTA with s.

More specifically, note that these four conditions imply (ml) but not (sm) or $\left(sm^{CU}\right)$ are EBAs between $S = sm$ and l in the subgame at the empty network. The former follows because $\langle (ml) \rangle$ is Pareto optimal for s and m and the fear of being an outsider in a Nash agreement deters any deviation by s or m. The latter follows because, given $\langle (ml) \rangle$ is Pareto dominant for l, m deviates from $a_m\left(\emptyset\right) \in \left\{s, s^{CU}\right\}$ to $a_m\left(\emptyset\right) = l$ and the unique Nash agreement (ml). Thus, l has a self enforcing deviation to (ml) from anything other than (ml). But, given g^{FT} is an equilibrium path of agreements, there must be a self enforcing deviation from (ml) to g^{FT}. Indeed, there is a self enforcing deviation from any action profile. Hence, given (ml) is an EBA between $S = sm$ and l, (ml) is an EBA in the subgame at the empty network and (ml, sl, sm) is an equilibrium path of agreements.

What does Proposition 2 say about the role of the tradeoff between the FTA flexibility benefit and the CU coordination benefit? It says this tradeoff plays little role when there is a small degree of asymmetry. When the CU coordination benefit dominates the FTA flexibility benefit, i.e. $\beta > \bar{\beta}_l^{Flex}\left(\alpha\right)$, l also favors expansion of $\left(ml^{CU}\right)$ to global free trade. Indeed, a direct move to global free trade is Pareto dominant for all players and, hence, the unique equilibrium path of agreements. Thus, the CU coordination benefit does not explain CU formation. Conversely, while a necessary condition for FTA formation in equilibrium is that the FTA flexibility dominates the CU coordination benefit, i.e. $\beta < \bar{\beta}_l^{Flex}\left(\alpha\right)$, it is not a sufficient condition (as demonstrated in the previous two paragraphs by reliance on $\langle (ml) \rangle \succ_m \langle \left(sm^{CU}\right) \rangle$). Therefore, I introduce a greater degree of asymmetry in the following section which will clearly bring out how the tradeoff between flexibility and coordination affects the equilibrium path of agreements.

5.2 "Moderate" degree of asymmetry

The area of the parameter space analyzed in this section, interpreted as a "moderate" degree of asymmetry, is summarized by three properties; the last two are preferences over paths of agreements. First, m and l hold a CU exclusion incentive. That is, m and l exclude s by refusing to expand their CU to global free trade. Second, l may prefer FTA rather than CU formation with m because the FTA flexibility benefit can dominate the CU coordination benefit. Third, m and l always prefer a CU with each other and the associated CU coordination benefit over an FTA with s and the associated FTA flexibility benefit. Condition 2 summarizes.

Condition 2 (Moderate degree of asymmetry)**.** Countries are asymmetric and i) $W_i\left(ml^{CU}\right) > W_i\left(g^{FT}\right)$ for $i = m, l$, ii) $\langle (ml) \rangle = (ml, sl, sm) \succ_l \left\langle \left(ml^{CU}\right) \right\rangle = \left(ml^{CU}\right)$ for some β, but iii) $\left\langle \left(ml^{CU}\right) \right\rangle = \left(ml^{CU}\right) \succ_m \langle (sm) \rangle = (sm, ml, sl)$ and $\left\langle \left(ml^{CU}\right) \right\rangle = \left(ml^{CU}\right) \succ_l \langle (sl) \rangle = (sl, ml, sm)$ for all β.

Condition 2 is captured in Figure 2 by area C2. The areas of the parameter space NA(a) and NA(b) are ignored to streamline the analysis. Their equilibrium characterization is tedious and no new insights emerge relative to those presented in the main text.[28]

Condition 2 violates Lemma 4 for two reasons. First, CU exclusion incentives. Indeed, $\left(ml^{CU}\right)$ is now Pareto dominant for m and, except for possibly $\langle (ml) \rangle$ or $\langle (sm) \rangle$, l as well. Second, despite the discrimination faced as an outsider, l may prefer to free ride on the external tariff liberalization created by the FTA between s and m rather than participate in global free trade. But, subject to these qualifications, Lemma 4 holds.

Lemma 6. *Assume Condition 2 holds. Lemma 4 holds except that i)* $W_i\left(ml^{CU}\right) > W_i\left(g^{FT}\right)$ *for* $i = m, l$ *and* $W_l\left(sl^{CU}\right) \lesseqgtr W_l\left(g^{FT}\right)$*, and ii)* $W_l(sm) \lesseqgtr W_l\left(g^{FT}\right)$.

Like earlier, the EBA at any hub–spoke network is an FTA between the spokes because it is mutually profitable. Thus, again, any hub–spoke network expands to global free trade.

The two issues that drive the EBA in subgames at FTA insider–outsider networks are the same as under a small degree asymmetry. The first issue is l's free riding incentives as an outsider. However, unlike under a small degree of asymmetry, s and m may remain insiders

[28]In area NA(a), $W_m\left(g^{FT}\right) > W_m\left(ml^{CU}\right)$ but $W_l\left(ml^{CU}\right) > W_l\left(g^{FT}\right)$ meaning that, unlike l, m wants to expand their CU to free trade. In area NA(b), $\langle (si) \rangle \succ_i \left\langle \left(ml^{CU}\right) \right\rangle$ for $i = m, l$ can occur meaning, for i, the flexibility benefit of an FTA with s can dominate the CU coordination benefit of $\left(ml^{CU}\right)$.

in the unique EBA. This arises when $W_l(sm) > W_l(g^{FT})$ because then $\bar{\beta}_l^{FT-O}(\alpha) > 1$: global free trade and elimination of the discrimination faced by l as an outsider does not outweigh l's incentive to free ride on the external tariff liberalization of s and m. Lemma 7 summarizes.

Lemma 7. *Consider the subgame at the FTA insider–outsider network* (sm). *No agreement is the unique EBA iff* $W_l(sm) > W_l(g^{FT})$.

The second important issue in subgames at FTA insider–outsider networks is the tradeoff that l faces between the CU coordination benefit and the FTA flexibility benefit. Given that l becomes the hub after forming an FTA with m but m and l hold a CU exclusion incentive, then l prefers to form an FTA over a CU with m if and only if

$$W_l(ml) + \beta W_l(lh) + \frac{\beta^2}{1-\beta} W_l(g^{FT}) \;>\; \frac{1}{1-\beta} W_l(ml^{CU}). \tag{5}$$

Rearranging (5) yields,

$$\underbrace{\beta\left[\left(W_l(lh) - W_l(ml^{CU})\right) + \frac{\beta}{1-\beta}\left(W_l(g^{FT}) - W_l(ml^{CU})\right)\right]}_{\text{FTA flexibility benefit}} > \underbrace{\left(W_l(ml^{CU}) - W_l(ml)\right)}_{\text{CU coordination benefit}}. \tag{6}$$

That is, l prefers an FTA over a CU with m iff the FTA flexibility benefit dominates the CU coordination benefit.

The decomposition of (5) into the FTA flexibility benefit and CU coordination differs from that under a small degree of asymmetry because (ml^{CU}) no longer expands to global free trade. Like before, the FTA flexibility benefit represents the additional continuation payoff of FTA formation over CU formation. But now the continuation payoff from CU formation is $\frac{1}{1-\beta}W_l(ml^{CU})$ rather than $\frac{1}{1-\beta}W_l(g^{FT})$ since CU formation no longer expands to global free trade. Thus, $W_l(lh) - W_l(ml^{CU}) > 0$ represents the immediate benefit l derives from sole preferential access in each spoke market as the hub relative to a CU with m. But, given the subsequent formation of a spoke–spoke FTA, $W_l(g^{FT}) - W_l(ml^{CU}) < 0$ reduces the value of the FTA flexibility benefit and may even make it negative.

The conditions under which the FTA flexibility benefit now dominate the CU coordination benefit are substantively different compared to a small degree of asymmetry. Rearranging (6), l prefers an FTA over a CU with m if and only if

$$\left[W_l(ml) - W_l(ml^{CU})\right] + \beta\left[W_l(lh) - W_l(ml)\right] + \beta^2\left[W_l(g^{FT}) - W_l(lh)\right] > 0. \tag{7}$$

Like under a small degree of asymmetry in (3), (7) fails when the discount factor is sufficiently small because of the immediate CU coordination benefit. However, unlike in (3), (7) also fails when the discount factor is sufficiently large because the part of the FTA flexibility that dominates is $W_l\left(g^{FT}\right) - W_l\left(ml^{CU}\right)$ which is negative because of l's CU exclusion incentive.[29] Thus, l prefers an FTA over a CU with m only when the discount factor is in an intermediate range $\left(\underline{\beta}_l^{Flex}(\alpha), \bar{\beta}_l^{Flex}(\alpha)\right)$.

Rolling back to the empty network, Proposition 3 characterizes the equilibrium path of agreements when the CU coordination benefit dominates the FTA flexibility benefit.

Proposition 3. *Assume Condition 2 holds and suppose the CU coordination benefit dominates the FTA flexibility benefit, $\beta \notin \left(\underline{\beta}_l^{Flex}(\alpha), \bar{\beta}_l^{Flex}(\alpha)\right)$. When l free rides in the subgame at (sm) and (sm) is Pareto dominant for l in the subgame at the empty network, the equilibrium paths of agreements are $\left(ml^{CU}\right)$ and (sm). Otherwise, $\left(ml^{CU}\right)$ is unique.*

Proposition 3 establishes a clear role for the trade off between flexibility and coordination: a sufficient condition for equilibrium CU formation is that the CU coordination benefit dominates the FTA flexibility benefit.

The intuition behind Proposition 3 is simple. Either $\langle\left(ml^{CU}\right)\rangle = \left(ml^{CU}\right)$ or $\langle(sm)\rangle$ is Pareto dominant for l but $\langle(sm)\rangle$ can only be Pareto dominant for l if l free rides on the FTA between s and m in the subgame at (sm). If $\left(ml^{CU}\right)$ is Pareto dominant for l then it is Pareto dominant for both m and l so Lemma 1 implies $\left(ml^{CU}\right)$ is the unique equilibrium path of agreements.

The logic is a little more involved for the case when free riding is Pareto dominant for l because there is a self enforcing deviation from any action profile in the subgame at the empty network. For $g \notin \left\{\left(ml^{CU}\right), (sm)\right\}$, m and l deviate to $\left(ml^{CU}\right)$. This is self enforcing because the fear of being a CU outsider in a Nash agreement deters l from attempting to subsequently deviate to $a_l(\emptyset) = \phi$. Moreover, m has a self enforcing deviation from $a_m(\emptyset) = s$ to $a_m(\emptyset) = l^{CU}$ and $\left(ml^{CU}\right)$ because $\left(ml^{CU}\right)$ is the only EBA between $S = sl$ and m when $a_m(\emptyset) = l^{CU}$. $\left(ml^{CU}\right)$ is the only such EBA because i) $\left(ml^{CU}\right)$ is Pareto dominant for m and, except for (sm), for l also, and ii) the fear of being a CU outsider in a Nash agreement deters any subsequent deviation by s or l. Additionally, s and l deviate from $\left(ml^{CU}\right)$ to $a_s(\emptyset) = m$ and $a_l(\emptyset) = \phi$ anticipating that m will accept the FTA with s. The deviation is self enforcing because, again, the fear of being a CU outsider deters any deviation by s or l. Thus, $\left(ml^{CU}\right)$ and (sm) are EBAs for the subgame at the empty network since they are EBAs between some S and $N \setminus S$. There are no other such EBAs because m or l deviate from $g \notin \left\{\left(ml^{CU}\right), (sm)\right\}$ to $a_m(\emptyset) = l^{CU}$ or $a_l(\emptyset) = m^{CU}$ and the unique Nash agreement $\left(ml^{CU}\right)$. Hence, $\left(ml^{CU}\right)$ and (sm) are the equilibrium paths of agreements.

[29] Indeed, (7) reduces to $W_l\left(g^{FT}\right) - W_l\left(ml^{CU}\right) < 0$ as $\beta \rightarrow 1$.

While formation of a single FTA is possible under Proposition 3 if there are multiple equilibria, it arises because of free riding incentives rather than the FTA flexibility benefit. However, Proposition 4 shows formation of *multiple* FTAs is tightly linked with the FTA flexibility benefit.

Proposition 4. *Assume Condition 2 holds and suppose the FTA flexibility benefit dominates the CU coordination benefit, $\beta \in \left(\underline{\beta}_l^{Flex}(\alpha), \bar{\beta}_l^{Flex}(\alpha)\right)$. When l does not free ride in the subgame at (sm) and m prefers a PTA with s over an FTA with l in the subgame at the empty network, the two equilibrium paths of agreements are (ml, sl, sm) and $\left(ml^{CU}\right)$. Otherwise, (ml, sl, sm) is unique.*

Together, Propositions 3 and 4 say the necessary and sufficient condition for (ml, sl, sm) to emerge as an equilibrium path of agreements is that the FTA flexibility benefit dominates the CU coordination benefit. Indeed, this is also the necessary and sufficient condition for the emergence of multiple FTAs in equilibrium. This formalizes the idea that the FTA flexibility benefit can play an important role in explaining the observed prevalence of FTAs relative to CUs.

While (ml, sl, sm) is always an equilibrium path of agreements, $\left(ml^{CU}\right)$ can be as well. Intuitively, the possibility hinges on whether m can credibly threaten to form a PTA with s rather than an FTA with l. If so, m's credible threat induces l to accept a CU with m rather than push for an FTA.

To formalize the intuition, note there is a self enforcing deviation from any action profile in the subgame at the empty network except, potentially, those supporting (ml). To see this note that $\langle\left(ml^{CU}\right)\rangle$ is Pareto dominant for m and l prefers $\langle\left(ml^{CU}\right)\rangle$ over $\langle g\rangle$ for any $g \notin \left\{(ml), (sm), \left(ml^{CU}\right)\right\}$. Thus, the only EBA between $S = sm$ or $S = ml$ and $N \setminus S$ is $\left(ml^{CU}\right)$. Moreover, the only EBAs between $S = sl$ and m are (1) (ml), (2) $\left(ml^{CU}\right)$ if m prefers a PTA with s over an FTA with l, and (3) (sm) if l free rides on the FTA between s and m and $\langle(sm)\rangle \succ_l \langle(ml)\rangle$.[30] These are EBAs between S and $N \setminus S$ because the fear of being a CU outsider in a Nash agreement deters any subsequent deviation by a country in the two country coalition. They are the only such EBAs because otherwise m or l deviate to $a_m(\emptyset) = l^{CU}$ or $a_l(\emptyset) = m^{CU}$ and the unique Nash agreement $\left(ml^{CU}\right)$. Thus, there are self enforcing deviations by i) l from $g \notin \left\{(ml), (sm), \left(ml^{CU}\right)\right\}$ to $\left(ml^{CU}\right)$, ii) either m or l from (sm) to, respectively, (ml) or $\left(ml^{CU}\right)$ and iii) $S = sl$ from $\left(ml^{CU}\right)$ to (ml). Moreover, there are only two possible self enforcing deviations from (ml): m to $\left(ml^{CU}\right)$ or $S = sl$ to (sm).

When l opts against free riding on the FTA between s and m, then $\langle(ml)\rangle \succ_l \langle(sm)\rangle$ and

[30] Note that $\langle(ml)\rangle \succ_m \langle g\rangle$ for $g \in \left\{(sm), \left(sm^{CU}\right)\right\}$ when l free rides in the subgame at (sm).

thus only m's deviation from (ml) warrants attention. Since the fear of being a CU outsider deters any subsequent deviation by s, m's deviation to $\left(ml^{CU}\right)$ is self enforcing if and only if the fear of a Nash agreement PTA between s and m deters the subsequent deviation by l from $a_l(\emptyset) = m^{CU}$ to $a_l(\emptyset) = m$. That is, m's deviation is self enforcing if and only if it can credibly threaten to form a PTA with s rather than an FTA with l. In this case, there is a self enforcing deviation from any action profile and the EBAs for the subgame at the empty network are the EBAs between some S and $N \setminus S$ which are (ml) and $\left(ml^{CU}\right)$. The two equilibrium paths of agreements are then (ml, sl, sm) and $\left(ml^{CU}\right)$. Otherwise, (ml) is the unique EBA of the subgame and the unique equilibrium path of agreements is (ml, sl, sm).

Interestingly, m's ability to credibly threaten l by forming a PTA with s depends crucially on whether l free rides on the FTA between s and m. l's free riding has this dramatic effect by affecting m's preferred type of PTA with s. m's FTA flexibility benefit derived from an FTA with s is quite large because it facilitates sole preferential access in l's market as the hub. As such, m prefers FTA formation over CU formation with s and possibly even an FTA with s rather than l. However, m does not become the hub when l free rides and, thus, m prefers to form an FTA with l rather than any PTA with s when l free rides. Hence, by free riding, l can effectively force m to accept an FTA rather than a CU because m cannot credibly threaten to form a PTA with s.

When l free rides, there is a second possible self enforcing deviation from (ml): s and l to (sm). Naturally, this requires $\langle(sm)\rangle \succ_{sl} \langle(ml)\rangle$. Since $\langle(sm)\rangle \succ_{sl} \langle(ml)\rangle$ requires not only that l free rides but also $W_l(sm) > W_l\left(g^{FT}\right)$, then $\langle(sm)\rangle \succ_l \langle(ml)\rangle$ when β is high because l prefers continuing to free ride as the outsider rather than end up in global free trade. However, s's incentives are opposite: $\langle(sm)\rangle \succ_s \langle(ml)\rangle$ when β is low because then s places great value on the immediate sole preferential access to m's market even though it never has preferential access to l's market as it eventually does under $\langle(ml)\rangle$. Indeed, $\langle(sm)\rangle \succ_l \langle(ml)\rangle$ implies $\langle(ml)\rangle \succ_s \langle(sm)\rangle$. Thus, there is no self enforcing deviation from (ml) when l free rides on the FTA between s and m. In turn, (ml) is the unique EBA in the subgame at the empty network and (ml, sl, sm) is the unique equilibrium path of agreements.

An interesting corollary emerges from Proposition 4 which highlights the importance of endogenously determining the type of PTA and also how FTAs play a subtle but important positive role.

Corollary 1 (Positive role of FTAs)**.** *Assume Condition 2 holds and suppose the the FTA flexibility benefit dominates the CU coordination benefit,* $\beta \in \left(\underline{\beta}_l^{Flex}(\alpha), \bar{\beta}_l^{Flex}(\alpha)\right)$*. Global free trade emerges in the presence of FTAs and CUs but does not emerge if CUs are the only type of PTA.*

If CUs are the only type of PTA, the CU exclusion incentive implies that, in equilibrium,

m and l form a CU and exclude s from expansion to global free trade. Indeed, this CU exclusion incentive underlies the result of Saggi et al. (2013) that CUs can prevent global free trade and the negative view of CUs in Missios et al. (2013). However, Corollary 1 says that FTA rather than CU formation occurs in equilibrium when the FTA flexibility benefit dominates the CU coordination benefit even if there is a CU exclusion incentive. Moreover, such FTA formation leads to global free trade. Thus, endogenously determining the type of PTA in a dynamic farsighted model can overturn the negative results of Saggi et al. (2013) and Missios et al. (2013). Additionally, the endogenous determination highlights that FTAs can play an important role in limiting the destructive nature of PTAs. Specifically, although FTAs emerge in equilibrium and lead to global free trade, global free trade would not emerge if CUs were the only type of PTA given the CU exclusion incentive of the two largest countries.

5.3 "Large" degree of asymmetry

The area of the parameter space analyzed in this section, interpreted as a "large" degree of asymmetry, is summarized by Condition 3 below and depicted in Figure 2 by areas C3(a) and C3(b). Note, $\bar{\alpha}_{ms} \approx 1.29$ is defined such that $W_m(sm) > W_m(\emptyset)$ if and only if $\alpha < \bar{\alpha}_{ms}$.

Condition 3 differs from Condition 2 in three respects. First, the CU coordination benefit that l and m receive from $\left(ml^{CU}\right)$ always dominates the FTA flexibility benefit. Second, from l's view, the only bilateral PTA that is necessarily welfare improving is a CU with m. To this end, the $\alpha_{ms} = \underline{\alpha}_{ms}(\alpha)$ line in Figure 2 is a contour curve where $W_l\left(ml^{CU}\right) > W_l(\emptyset)$ above the contour curve. Third, only imposing $W_l\left(ml^{CU}\right) > W_l(\emptyset)$ allows the possibility that $W_l(\emptyset) > W_l\left(g^{FT}\right)$. Indeed, the $\alpha_{ms} = \bar{\alpha}_{ls}^{MFN}(\alpha)$ line in Figure 2 is a contour curve where $W_l\left(g^{FT}\right) > W_l(\emptyset)$ above the contour curve.

Condition 3 (Large degree of asymmetry)**.** Countries are asymmetric and i) $\left\langle\left(ml^{CU}\right)\right\rangle = \left(ml^{CU}\right) \succ_l \langle(ml)\rangle = (ml, sl, sm)$ and $\left\langle\left(ml^{CU}\right)\right\rangle = \left(ml^{CU}\right) \succ_m \langle(sm)\rangle = (sm, ml, sl)$ for all β, ii) $W_l\left(ml^{CU}\right) > W_l(\emptyset)$ and $W_m\left(ml^{CU}\right) > W_m\left(g^{FT}\right)$, and iii) $\alpha_{ms} < \bar{\alpha}_{ms}$ and $\alpha_{ls} < \bar{\alpha}_{ms}^2$.

Subject to the caveats embodied in Condition 3, Lemma 6 still holds.

Lemma 8. *Assume Condition 3 holds. Lemma 6 holds except that* $W_l(\emptyset) \gtrless W_l\left(g^{FT}\right)$ *and* $W_l(g) \gtrless W_l(\emptyset)$ *for* $g \in \left\{(sl), \left(sl^{CU}\right), (ml)\right\}$.

Proposition 5 characterizes the equilibrium path of agreements in a simple manner.

Proposition 5. *Assume Condition 3 holds. If* $\langle(sm)\rangle \succ_l \left\langle\left(ml^{CU}\right)\right\rangle = \left(ml^{CU}\right)$, *the two equilibrium paths of agreements are* $\left(ml^{CU}\right)$ *and* (sm). *Otherwise,* $\left(ml^{CU}\right)$ *is unique.*

The simplicity of the equilibrium structure stems from the fact that $\langle (ml^{CU}) \rangle = (ml^{CU})$ is Pareto dominant for m and, absent $\langle (sm) \rangle$, for l as well. Lemma 1 implies (ml^{CU}) is the unique equilibrium path of agreements when it is Pareto dominant for m and l. This happens if and only if $W_l(ml^{CU}) > W_l(sm)$.[31] Moreover, since $W_l(sm) > W_l(ml^{CU})$ requires $W_l(sm) > W_l(g^{FT})$, $\langle (sm) \rangle$ can be Pareto dominant for l only if it free rides on the FTA between s and m (see Lemma 7) in the subgame at (sm).

When l free rides on the FTA between s and m in the subgame at (sm) and (sm) is Pareto dominant for l in the subgame at the empty network, there is a self enforcing deviation from any action profile in the subgame at the empty network. Following the logic of Proposition 4 in the analogous case when l free rides, (ml^{CU}) is the EBA between $S = sm$ or $S = ml$ and $N \setminus S$ while (ml^{CU}) and (sm) are the EBAs between $S = sl$ and m. Thus, there are self enforcing deviations by m from $g \neq (ml^{CU})$ to (ml^{CU}) and $S = sl$ from (ml^{CU}) to (sm). The fear of being a CU outsider in a Nash network deters any subsequent deviation by a country in the two country coalition. Thus, the EBAs for the subgame are the EBAs between some S and $N \setminus S$: (ml^{CU}) and (sm). These are also the equilibrium paths of agreements. In any case, Proposition 5 highlights that global free trade is never attained.

6 Application to building bloc–stumbling bloc issue

This section compares the equilibrium outcomes of two games: the one of previous sections where countries chose between formation of PTAs and MFN agreements and one where PTAs do not exist. Recently, Saggi and Yildiz (2010, 2011) and Saggi et al. (2013) have used this equilibrium comparison approach (although each of those papers considers either FTAs or CUs as the only type of PTA) to analyze the notion of whether PTAs are necessary for global free trade or prevent global free trade. That is, in the famous terminology of Bhagwati (1991, 1993), are PTAs building blocs or stumbling blocs to global free trade?

Undertaking this analysis first requires solving the equilibrium outcome in the absence of PTAs. The following proposition follows directly from Lemma 1.

Proposition 6. *In the absence of PTAs, the unique equilibrium path of agreements is a direct move to global free trade if* $W_l(g^{FT}) \geq W_l(\emptyset)$ *and no agreement otherwise.*

Thus, in the absence of PTAs, a direct move to global free trade emerges as the unique equilibrium path of agreements in area C3(a) but not C3(b) given the definition of $\bar{\alpha}_{ls}^{MFN}(\alpha)$ (see discussion above Lemma 8).

[31] $W_l(ml^{CU}) > W_l(sm)$ implies $W_l(ml^{CU}) > \max\{W_l(sm), W_l(mh), W_l(g^{FT})\}$ and thus $\langle (ml^{CU}) \rangle = (ml^{CU}) \succ_l \langle (sm) \rangle$. $W_l(ml^{CU}) < W_l(sm)$ implies $W_l(sm) > W_l(g^{FT})$ which, in turn, implies no agreement is the EBA in the subgame at (sm) and $\langle (sm) \rangle = (sm) \succ_l \langle (ml^{CU}) \rangle = (ml^{CU})$.

Proposition 6 says a direct move to global free trade is the unique equilibrium in the absence of PTAs unless the large country blocks global free trade. That is, market access to the largest country is valuable enough to the other countries that they never block global free trade. In contrast, the market access gained in other countries by the largest country may not adequately compensate for the domestic market access given up. Intuitively, the large country blocks global free trade when it views the world market as too small.

Corollary 2, following directly from Propositions 1, 2 and 6, summarizes the irrelevance of PTAs in attaining global free trade under symmetry or a small degree of asymmetry.

Corollary 2 (Irrelevance of PTAs)**.** *Under symmetry or a small degree of asymmetry (i.e. Condition 1), global free trade is achieved regardless of whether PTAs exist or not.*

The only difference between the equilibrium outcomes in the presence of PTAs under symmetry versus a small degree of asymmetry is that global free trade obtains immediately under the former but may be attained via a path of PTAs under the latter. Thus, in either of these cases, PTAs neither prevent nor are necessary for global free trade.

However, Corollary 3, which follows directly from Propositions 3 and 6, shows the benign effects of PTAs disappear under a moderate degree of asymmetry.

Corollary 3 (Destructive role of PTAs)**.** *Consider a moderate degree of asymmetry (i.e. Condition 2 holds) and suppose the CU coordination benefit dominates the FTA flexibility benefit,* $\beta \notin \left(\underline{\beta}_l^{Flex}, \bar{\beta}_l^{Flex}\right)$*. PTAs prevent global free trade since global free trade is not attained in the presence of PTAs but is attained in their absence.*

Corollary 3 emphasizes the destructive role that PTAs play in the model. Under Condition 2, the degree of asymmetry is moderate enough that l does not veto a direct move to global free trade in the absence of PTAs. That is, the world market is big enough that the market access received by l compensates it for the domestic market access given up and so it will participate in global free trade when MFN liberalization is the only form of liberalization. However, Proposition 3 states that global free trade is not attained when the CU coordination benefit dominates the FTA flexibility benefit. That is, even though global free trade is attained in the absence of PTAs, countries choose to form PTAs in equilibrium when given the opportunity and, as a result, global free trade does not emerge. The basic reason behind the destructive role of PTAs is the CU exclusion incentive held by m and l: m and l refuse to expand their CU to global free trade.

The destructive role of PTAs described in Corollary 3 is not in itself surprising. For example, following the same equilibrium comparison approach, Saggi et al. (2013) find PTAs may prevent global free trade because of the CU exclusion incentive. However, they do not endogenize the type of PTA. CUs are the only type of PTA in their model. Thus, Corollary 3 goes further than Saggi et al. (2013). Specifically, it says the tradeoff between the CU coordination benefit and the FTA flexibility benefit, which is a fundamentally dynamic issue, drives whether PTAs prevent global free trade even in the presence of CU exclusion incentives.

Endogenizing the choice between FTAs and CUs has strong implications. Once the type of PTA is endogenous, PTAs need not prevent global free trade even in the presence of the CU exclusion incentives. Corollary 4 follows directly from Propositions 4 and 6.

Corollary 4 (Free riding and positive role of FTAs.). *Consider a moderate degree of asymmetry (i.e. Condition 2 holds) and suppose the FTA flexibility benefit dominates the CU coordination benefit ,* $\beta \in \left(\underline{\beta}_l^{Flex}(\alpha), \bar{\beta}_l^{Flex}(\alpha)\right)$*. A path of FTAs leading to global free trade is an equilibrium path of agreements and unique if* l *free rides on the off equilibrium path FTA between* s *and* m*.*

While previous papers (e.g. Saggi and Yildiz (2010)) have found similar benign effects of FTAs, Corollary 4 goes further than existing results for two reasons. First, previous results were derived in frameworks that did not endogenize the choice between FTAs and CUs. This is an important qualification because Corollary 4 says PTAs can play a benign role even in the presence of CU exclusion incentives which drive the negative view of PTAs in other models such as Saggi et al. (2013) and Missios et al. (2013). Thus, recognizing the FTA flexibility benefit can overturn existing results in the literature. Importantly, this insight only emerges because the model is dynamic.

Second, and surprisingly, Corollary 4 explains how FTA free riding incentives actually increase the extent that FTAs rather than CUs emerge in equilibrium and thus the extent that PTAs play a benign role. The basic intuition is simple. By free riding on the FTA between s and m, l can effectively force m to accept an FTA rather than a CU. Underlying this logic is that m cannot credibly threaten to form a PTA with s since l's free riding destroys the FTA flexibility benefit that m derives from FTA formation with s and, thus, m prefers FTA formation with l over any PTA with s.

The above corollaries establish that whether PTAs prevent global free trade or not depends crucially on the tradeoff between the FTA flexibility benefit and the CU coordination benefit. However, the possibility that PTAs could be necessary for global free trade was implicitly removed because, by Proposition 6, global free trade fails to arise in the absence of PTAs only if there is a large degree of asymmetry. Corollary 5, following directly from Propositions 5 and 6, addresses whether PTAs can be necessary for global free trade.

Corollary 5 (PTAs play no constructive role)**.** *When global free trade is not attained in the absence of PTAs, the coordination benefit of CUs dominates the flexibility benefit of FTAs and PTAs do not lead to global free trade. Thus, PTAs are never necessary for global free trade.*

Corollary 5 extends Corollary 3 by establishing that, in addition to the potential destructive role of PTAs, PTAs never play a constructive role in attaining global free trade. In this sense, the paper provides a very negative view of PTAs. As discussed in the introduction, the result that PTAs are never necessary for global free trade contrasts with recent results that emphasize PTAs may be necessary for global free trade (e.g. Ornelas (2007), Saggi and Yildiz (2010) and Saggi et al. (2013)).

For Saggi and Yildiz (2010) and Saggi et al. (2013), a key ingredient establishing that PTAs can be necessary for global free trade is that, in the absence of PTAs, countries may want to free ride on joint MFN tariff liberalization by other countries. When a country deviates from global free trade and anticipates MFN tariff reductions by the other two countries (i.e. a two country MFN agreement), global free trade may not emerge. Thus, the presence of two country MFN agreements strengthen the possibility that PTAs play a constructive role.

While this paper has not considered two country MFN agreements, Proposition 7 below shows this does not alter the result that PTAs are never necessary for global free trade. Incorporating two country MFN agreements requires modifying the action space. In particular, $A_i(\emptyset) = \left\{\phi, FT, j^M, k^M\right\}$ where j^M and k^M indicates i announces it wants to form a two country MFN agreement with j or k respectively. Moreover, denoting a two country MFN agreement by $\left(ij^M\right)$, $A_\iota\left(ij^M\right) = \{\phi, FT\}$ for any $\iota \in N$. As throughout the paper, an agreement forms if and only if all members of the proposed agreement announce in favor.[32]

Proposition 7. *In the absence of PTAs, suppose countries can form two or three country MFN agreements. A direct move to global free trade is the unique equilibrium path of agreements if $W_l\left(g^{FT}\right) > W_l(\emptyset)$.*

The key intuition is that even though l may prefer to free ride on MFN tariff liberalization between s and m rather than participate in global free trade, such liberalization is not an equilibrium outcome if l backs out of global free trade. That is, l's deviation from $a_l(\emptyset) = FT$ to $a_l(\emptyset) = \phi$ is not self enforcing because $\left(sm^M\right)$ is not an EBA between $S = sm$ and l. The reason is that m will subsequently back out of MFN negotiations with s and announce it wants to form a two country MFN agreement with l. At that stage, l's

[32]Tariffs of a two country MFN agreement are chosen to maximize joint member welfare (like Saggi and Yildiz (2010) and Saggi et al. (2013)) subject to the constraint that each members welfare rises.

best response is to accept given the unique Nash agreement is $\left(ml^M\right)$. Thus, given this anticipation, l will not deviate from global free trade.

7 Conclusion

Since the early 1990s, the number of PTAs has expanded exponentially. However, while some influential PTAs are CUs, the vast majority of PTAs are FTAs. Indeed, 164 out of the 169 PTAs notified to the WTO under GATT Article XXIV since 2000 are FTAs.[33] This is surprising given that CU members coordinate on common external tariffs. Indeed, dating back to Kennan and Riezman (1990), the literature has long recognized this coordination benefit of CUs. To this end, Melatos and Woodland (2007, p.904) state that "... the apparent inconsistency between the observed popularity of free trade areas [FTAs] and the theoretical primacy of customs unions..." remains an unresolved issue and Facchini et al. (2012, p.136) state "... the existing literature has indicated that CUs are... the optimal form of preferential agreements [for members].".

I propose a novel idea underpinning this observed prevalence of FTAs relative to CUs: FTAs posses a dynamic flexibility benefit because, unlike individual CU members, individual FTA members have the flexibility to form their own subsequent agreements. Indeed, as described in the introduction, the idea of the FTA flexibility benefit matches well with recent discussions in the media regarding Uruguay's inability to form an FTA with the US given its membership in the MERCOSUR CU and drawbacks faced by the UK and Turkey due to, respectively, being a member of the EU CU or having a CU with the EU.

Whether FTAs emerge in equilibrium, and whether there is a real trade off between the FTA flexibility benefit and the CU coordination benefit, depends in a meaningful way on the degree of market size asymmetry. With a small degree of asymmetry, a direct move to global free trade emerges and there is no real trade off. With a large degree of market size asymmetry, a permanent CU between the largest countries emerges because the CU coordination benefit dominates the FTA flexibility benefit. But with a moderate degree of asymmetry, the trade off between the FTA flexibility and CU coordination benefits is fundamental to the equilibrium type of PTA. A sufficient condition for equilibrium CU formation is that the CU coordination benefit dominates the FTA flexibility benefit. Conversely, a necessary and sufficient condition for multiple FTAs in equilibrium is that the FTA flexibility benefit dominates the CU coordination benefit. Thus, the FTA flexibility benefit is crucial to explaining the prevalence of FTAs relative to CUs in the model.

Because of the endogenous choice between FTAs and CUs in a setting where there is a

[33] `http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx`

trade off between the FTA flexibility and CU coordination benefits, the model also produces insights on the long standing building bloc–stumbling bloc issue. On one hand, FTAs may emerge in equilibrium and lead to global free trade even in cases (such as Saggi et al. (2013) and Missios et al. (2013)) where CUs would prevent global free trade if CUs were the only type of PTA. This possibility arises because of the FTA flexibility benefit and reflects a subtle but potentially important way in which FTAs mitigate the destructive role played by CU members excluding nonmembers from future expansion. Surprisingly, the role of the FTA flexibility benefit is magnified when there are off the equilibrium path FTA free riding incentives. On the other hand, when the CU coordination benefit makes the CU exclusion incentive strong enough, CU rather than FTA formation may emerge in equilibrium and prevent global free trade even though PTAs would not play this destructive role if FTAs were the only type of PTA. Indeed, this observation helps explain the unusually strong result that PTAs can prevent global free trade but are never necessary for global free trade.

Appendix

A One period welfare differences across networks

Let $\eta_i(g)$ denote the number of, and let g_i denote the set of, countries j such that $\tau_{ji}(g) = 0$. Then,

$$CS_i(g) = \frac{1}{32}\left[3\alpha_i + (\eta_i(g) - 3) \cdot \bar{\tau}_i(g)\right]^2,$$

$$\pi_i(g) = \frac{1}{16}\left[\sum_{j \in g_i}\left[\alpha_j + (3 - \eta_j(g)) \cdot \bar{\tau}_j(g)\right]^2 + \sum_{j \notin g_i}\left[\alpha_j - (1 + \eta_j(g)) \cdot \bar{\tau}_j(g)\right]^2\right],$$

and

$$TR_i(g) = \frac{1}{4}(3 - \eta_i(g)) \cdot \bar{\tau}_i(g) \cdot \left[\alpha_i - (1 + \eta_i(g)) \cdot \bar{\tau}_i(g)\right].$$

Denote a variable without its fraction of proportionality by adding a tilde. For example, $\tilde{CS}(g) \equiv 32CS(g)$ and $\tilde{q}^*_{ij}(g) \equiv 4q^*_{ij}(g)$. Then, multiplying each element of $W_i(g)$ by 32,

$$W_i(g) \propto \hat{W}(g) \equiv \tilde{CS}_i(g) + 2\tilde{\pi}_i(g) + 8\tilde{TR}_i(g)$$

and $(W_i(g') - W_i(g)) \propto \left(\hat{W}_i(g') - \hat{W}_i(g)\right)$ for any g', g. Of course, this proportionality representation is also true for linear combinations of differences.

Let $g + ij$ denote the network that adds a PTA between i and j to the network g. Also define $\bar{\tau}_{i0} \equiv \bar{\tau}_i(g)$, $\bar{\tau}_{i1} \equiv \bar{\tau}_i(g + ij)$, $\Delta\bar{\tau}_i \equiv \bar{\tau}_{i1} - \bar{\tau}_{i0}$, $\eta_{i0} \equiv \eta_i(g)$ and $\eta_{i1} \equiv \eta_i(g + ij)$. Then, letting $b_i(\cdot) \equiv (\eta_{i0} - 3)\Delta\bar{\tau}_i + \bar{\tau}_{i1}$,

$$\Delta\tilde{CS}_i(g+ij) = b_i(\cdot)\left[b_i(\cdot) + 2\tilde{Q}_i(g)\right],$$

$$\Delta\tilde{\pi}_{ii}(g+ij) = b_i(\cdot)\left[b_i(\cdot) - 2\tilde{q}^*_{ii}(g)\right],$$

$$\Delta\tilde{\pi}_{ij}(g+ij) = b_j(\cdot)\left[b_j(\cdot) - 2\tilde{q}^*_{ij}(g)\right] + 8\bar{\tau}_{j0}\left[2\bar{\tau}_{j0} + \tilde{q}^*_{ij}(g) - b_j(\cdot)\right],$$

and

$$\Delta\tilde{TR}_i(g+ij) = -b_i(\cdot)\tilde{q}^*_{ji}(g) - (2-\eta_{i0})\bar{\tau}_{i1}\left[b_i(\cdot) + 4\Delta\bar{\tau}_i\right].$$

B Proofs

As noted at the end of Section 3.3, I now outline notation that helps facilitate exposition of the proofs. Let $G(P,g)$ be the set of networks $g' = g + \ell$ in the subgame at network g where ℓ is an EBA given the coalition structure P.[34] Specifically, $P = P^* \equiv \{\{i\},\{j\},\{k\}\}$ is the singletons coalition structure, $P = P_S \equiv \{\{S\},\{N \setminus S\}\}$ is the coalition structure with S and $N \setminus S$ as the two coalitions and $P = N$ is the coalition structure with the grand coalition. Thus, for example, $G(P_S,g)$ is the set of networks $g' = g + \ell$ in the subgame at network g where ℓ is an EBA between S and $N \setminus S$. Similarly define $\gamma(P,g)$ as the set of networks $g' = g + \ell$ in the subgame at network g where ℓ is a Nash agreement given the coalition structure P. For example, $\gamma(P_S,g)$ is the set of networks $g' = g + \ell$ in the subgame at network g where ℓ is a Nash agreement between S and $N \setminus S$ (i.e. the action profile inducing ℓ is a Nash equilibrium between S and $N \setminus S$).

Proof of Lemma 1: Parts i) and ii) follow from Proposition 4 of Lake (2013). Part iii) is Lemma 1 from Lake (2013).■

Proof of Lemma 2: The proof relies on derivations in Appendix A. For part i) notice that $\hat{W}_i(jh) - \hat{W}_i(jk) = -1.3717\alpha_i^2 + 1.347\alpha_j^2 < 0$ but $\hat{W}_i(ij) - \hat{W}_i(\emptyset) = -1.3717\alpha_i^2 + 2.2922\alpha_j^2 > 0$, $\hat{W}_i(ih) - \hat{W}_i(ik) = -.4283\alpha_i^2 + 2.2922\alpha_j^2 > 0$, $\hat{W}_i(g^{FT}) - \hat{W}_i(kh) = -.4283\alpha_i^2 + 1.347\alpha_j^2 > 0$, and $\hat{W}_i(ij^{CU}) - \hat{W}_i(\emptyset) = -1.3398\alpha_i^2 + 1.8777\alpha_j^2 + .658\alpha_i\alpha_j > 0$. For part ii), $\hat{W}_i(ih) - \hat{W}_i(g^{FT}) = .6122(\alpha_j^2 + \alpha_k^2) > 0$, $\hat{W}_i(g^{FT}) - \hat{W}_i(ij) = -.4283\alpha_i^2 - .6122\alpha_j^2 + 1.68\alpha_k^2 > 0$, $\hat{W}_i(g^{FT}) - \hat{W}_i(ij^{CU}) = -.4602\alpha_i^2 - .1977\alpha_j^2 - .658\alpha_i\alpha_j + 1.68\alpha_k^2 > 0$ and $\hat{W}_i(g^{FT}) - \hat{W}_i(jk) = -1.8\alpha_i^2 + 1.347(\alpha_j^2 + \alpha_k^2) > 0$. Note, $\hat{W}_i(g^{FT}) > \hat{W}_i(jk)$ and $\hat{W}_i(jk) > \hat{W}_i(\emptyset) > \hat{W}_i(jk^{CU})$ implies $\hat{W}_i(g^{FT}) > \hat{W}_i(\emptyset)$ and $\hat{W}_i(g^{FT}) > \hat{W}_i(jk^{CU})$. For part iii), $\hat{W}_i(ij^{CU}) - \hat{W}_i(ij) = .0319\alpha_i^2 - .4145\alpha_j^2 + .658\alpha_i\alpha_j > 0$. For future reference, when the

[34] $g' = g + \ell$ is standard network notation indicating that the agreement ℓ is added to the network g.

CU external tariff constraint binds then $\hat{W}_i\left(ij^{CU}\right) - \hat{W}_i(\emptyset) = -1.8\alpha_i^2 + 1.17\alpha_j^2 + 1.8\alpha_i\alpha_j$ if $\alpha_i > \alpha_j$ and $\hat{W}_i\left(ij^{CU}\right) - \hat{W}_i(\emptyset) = -1.71\alpha_i^2 + 1.68\alpha_j^2 + 1.2\alpha_i\alpha_j$ if $\alpha_i < \alpha_j$. ■

Proof of Lemma 3: Note that in any subgame at a hub–spoke network $g = (ih)$, Lemmas 2 and 1 imply that $G(N,(ih)) = g^{FT}$. Thus, $\langle (ih)\rangle \succ_i \langle (ij)\rangle$ since $W_i(ih) > W_i\left(g^{FT}\right) > W_i(ij)$.

To begin, first let $\langle (ih)\rangle \succ_k \langle (ij)\rangle$. Thus, $(ij) \notin \gamma(P_{\iota k},(ij))$ for $\iota = i,j$ because $\langle (\iota h)\rangle \succ_{\iota k} \langle (ij)\rangle$. Also, given Lemma 2, $\gamma(P^*,(ij)) = \left\{(ij),(ih),(jh),g^{FT}\right\}$ where $(ij) \in \gamma(P^*,(ij))$ implies $a_i(ij) \neq k$ and $a_j(ij) \neq k$. Moreover, $g^{FT} \in G(P_{ik},(ij))$ and, by symmetry, $g^{FT} \in G(P_{jk},(ij))$. This follows because i) $(jh) \in \gamma(P^*,(ij))$ deters i's deviation from g^{FT} to (ih), ii) $\left\langle g^{FT}\right\rangle$ is Pareto dominant for k and iii) $W_j\left(g^{FT}\right) > W_j(ik)$. Hence, $G(N,(ij)) \subseteq g^{FT}$ because k and some $\iota \in \{i,j\}$ deviates from any $g \in \{(ij),(ih),(jh)\}$ to $g^{FT} \in G(P_{\iota k},(ij))$.

To establish $g^{FT} \in G(N,(ij))$, restrict attention to unilateral deviations by i given $\left\langle g^{FT}\right\rangle$ is Pareto dominant for k and, by Lemma 2, i could only prefer $\langle ih\rangle$ over $\left\langle g^{FT}\right\rangle$. Thus, given $(ij) \notin \gamma(P_{jk},(ij))$, i has a self enforcing deviation from g^{FT} to (ih) iff $(ih) \in G(P_{jk},(ij))$ and $(jh) \notin G(P_{jk},(ij))$. $(ih) \in G(P_{jk},(ij))$ follows because i) $(ih) \in \gamma(P_{jk},(ij))$, ii) $(ih) \in \gamma(P^*,(ij))$ deters any deviation by j, and iii) $(ij) \in \gamma(P^*,(ij))$ deters k's deviation to g^{FT}. Moreover, $(jh) \notin G(P_{jk},(ij))$ iff $\langle (ih)\rangle \succ_k \langle (jh)\rangle$ given that k cannot unilaterally deviate to g^{FT} and $\langle (\iota h)\rangle$ is Pareto dominant for ι. But, $\langle (ih)\rangle \succ_k \langle (jh)\rangle$ iff $\alpha_i > \alpha_j$ which fails under symmetry. Thus, $(jh) \in G(P_{jk},(ij))$ deters i's deviation from g^{FT}. Therefore, $g^{FT} = G(N,(ij))$.

Second, let $\langle (ij)\rangle \succ_k \langle (ih)\rangle$ where $\alpha_i \geq \alpha_j$. Thus, $\gamma(P^*,(ij)) = \left\{(ij),g^{FT}\right\}$. $g^{FT} = G(N,(ij))$ follows from (a) $g^{FT} \in G(P_S,(ij))$ for $S \subset N$ but (b) $(\iota h) \notin G(P_S,(ij))$ for $\iota \in \{i,j\}$ and $S \subset N$. Specifically, (a) implies that some $S' = \iota k$, $\iota \in \{i,j\}$, deviates from $g \neq g^{FT}$ to $g^{FT} \in G(P_{S'},(ij))$ while (b) and part ii) of Lemma 2 imply $g^{FT} \in G(N,(ij))$ because there is no deviation from g^{FT} to any $g \in G(P_S,(ij))$. To see (a) holds, note $\left\langle g^{FT}\right\rangle$ is Pareto dominant for k and $(ij) \in \gamma(P^*,(ij))$ deters the deviation by $\iota = i,j$ from $a_\iota(ij) = FT$ to $a_\iota(ij) = k$. (b) holds because, for $\iota = i,j$, $a_k(ij) = \phi$ is k's best response to $a_\iota(ij) = k$ which implies $(\iota h) \notin \gamma(P_{ij},(ij))$ while $(\iota h) \notin G(P_{\iota k},(ij))$ because k has a self enforcing deviation from $a_k(ij) = \iota$ to $a_k(ij) = \phi$ given $\gamma(P^*,(ij)) = (ij)$ when $a_k(ij) = \phi$.■

Proof of Lemma 4: This follows using the one period payoff equations presented in Lemma 2; in particular note they are increasing in α_j.■

Proof of Lemma 5: Since Lemma 4 implies $\left\langle g^{FT}\right\rangle$ is Pareto dominant for k, the proof closely follows that of Lemma 3. For the case of $\langle (ij)\rangle \succ_k \langle (ih)\rangle$, which happens iff $k = l$ and $\beta < \bar{\beta}_l^{FT-O}(\alpha)$ given Condition 1, the proof is identical.

When$\langle (ih)\rangle \succ_k \langle (ij)\rangle$ then, like Lemma 3, $g^{FT} \in G(P_{jk},(ij))$. Thus, $G(N,(ij)) \subseteq$

$\{g^{FT},(jh)\}$ because $S=jk$ deviate from $g\in\{(ij),(ih)\}$ to $g^{FT}\in G(P_{jk},(ij))$. Moreover, like Lemma 3, $(ih)\in G(P_{jk},(ij))$ while $(jh)\notin G(P_{jk},(ij))$ iff $\langle(ih)\rangle\succ_k\langle(jh)\rangle$. However, unlike Lemma 3, $\langle(ih)\rangle\succ_k\langle(jh)\rangle$. Thus, i deviates from $g\in\{g^{FT},(jh)\}$ to $(ih)\in G(P_{jk},(ij))$. Hence, $G(N,(ij))$ is empty. $G^{EBA}=\underset{S\subset N}{\cup}G(P_S,(ij))=\{g^{FT},(ih)\}\equiv\tilde{G}$ then follows because $\gamma(P^*,(ij))=(ih)$ when $a_i(ij)=k$ or $a_k(ij)=i$ so, for $g\notin\tilde{G}$, i or k deviate from $g\in\gamma(P_S,(ij))$, $S\subset N$, to $(ih)=\gamma(P^*,(ij))$.■

Proof of Proposition 2: In any subgame at a hub–spoke network $g=(ih)$, Lemmas 4 and 1 imply $G(N,(ih))=g^{FT}$. Similarly, in any subgame at the CU insider–outsider network $g=(ij^{CU})$, Lemmas 4 and 1 imply $G(N,(ij^{CU}))=g^{FT}$. In any subgame at an FTA insider–outsider network $g=(ij)$, Lemma 5 implies $G(N,(ij))=g^{FT}$ if $k=l$ and $\beta<\bar{\beta}_l^{FT-O}(\alpha)$ but $G(N,(ij))=\{g^{FT},(ih)\}$ otherwise where $\alpha_i>\alpha_j$.

Now consider the subgame at the empty network $g=\emptyset$. For the conditions under which g^{FT} is unique, note that Lemma 4 implies a sufficient condition under which $\langle g^{FT}\rangle$ is Pareto dominant for N is $(W_i(g^{FT})-W_i(ij))+(W_i(g^{FT})-W_i(ih))=-1.8(\alpha_i)^2+1.0678(\alpha_j^2+\alpha_k^2)>0$. Hence, g^{FT} is the unique equilibrium path of agreements if $\alpha_{ls}<\sqrt{.59(\alpha_m^2+\alpha_s^2)}$.

Non uniqueness of g^{FT} follows by example. Let $\alpha_{ls}=\frac{7}{6}$, $\alpha_{ms}=\frac{101}{100}$ and $\beta=\frac{1}{2}$. Then, $\bar{\beta}_l^{FT-H}(\alpha)=.382$ and $\bar{\beta}_l^{FT-O}(\alpha)=.645$ meaning $\langle(ml)\rangle$ is Pareto dominant for l and $G(N,(sm))=g^{FT}$. Moreover, $\langle(ml)\rangle\succ_m\langle(sm^{CU})\rangle$ because $(W_m(ml)-W_m(sm^{CU}))+\beta(W_m(lh)-W_m(g^{FT}))>0$ reduces to $\beta<.6$. Thus, $\langle(ml)\rangle\succ_m\langle g\rangle$ for $g\notin\{(ml^{CU}),g^{FT}\}$. Additionally, $\langle(sl)\rangle\succ_l\langle(ml^{CU})\rangle$ because $(W_l(sl)-W_l(ml^{CU}))+\beta(W_l(lh)-W_l(g^{FT}))$ reduces to $\beta>.357$ which implies $(sl)\in\gamma(P^*,\emptyset)$ when $a_m(\emptyset)=l^{CU}$. Finally, $\langle(ml)\rangle\succ_s\langle(ml^{CU})\rangle$ because $(W_s(ml)-W_s(ml^{CU}))+\beta(W_s(lh)-W_s(g^{FT}))$ reduces to $\beta<1.396$.

These observations imply $(ml)\in G(P_{sm},\emptyset)$ because $(ml)\in\gamma(P_{sm},\emptyset)$ while $(sl)\in\gamma(P^*,\emptyset)$ deters any deviation by m and $(ml)\in\gamma(P^*,\emptyset)$ deters any deviation by s. Moreover, $G(P_{sm},\emptyset)\subseteq\{(ml),(ml^{CU}),g^{FT}\}\equiv G_1$ because m has a self enforcing deviation from $g\notin G_1$ to $a_m(\emptyset)=l$ since $(ml)=\gamma(P^*,\emptyset)$ when $a_m(\emptyset)=l$. Thus, $g^{FT}\notin G(N,\emptyset)$ since l has a self enforcing deviation from g^{FT} to $(ml)\in G(P_{sm},\emptyset)$. But, by part ii), g^{FT} is an equilibrium path of agreements. Hence, $G(N,\emptyset)$ is empty and g^{FT} is not the unique equilibrium path of agreements since the EBAs are $G^{EBA}\equiv\underset{S\subset N}{\cup}G(P_S,\emptyset)\supseteq(ml)$.

To see the eventual attainment of global free trade in any equilibrium path of agreements, note that this is true if no agreement, i.e. $\emptyset$, is not an EBA of the subgame. This is established by showing that i) $\emptyset\notin G(N,\emptyset)$ and $\emptyset\notin G^{EBA}\equiv\underset{S\subset N}{\cup}G(P_S,\emptyset)$ but ii) $g^{FT}\in G(P_{sm},\emptyset)$. First, $W_i(g^{FT})>W_i(\emptyset)$ for all i implies $\emptyset\notin G(N,\emptyset)$ and $W_i(g^{FT})>W_i(ij^{CU})>W_i(\emptyset)$ for any i,j implies $\emptyset\notin\gamma(P_{ij},\emptyset)$ for any i,j. Second, $g^{FT}\in\gamma(P_{sm},\emptyset)$ since i) $\langle g^{FT}\rangle$ is Pareto dominant for s and ii) Lemma 4 implies $\langle g^{FT}\rangle$ is Pareto dominant for m unless

$\langle (sm) \rangle \succ_m \langle g^{FT} \rangle$ in which case $(sl) \in \gamma(P^*, \emptyset)$ deters the m's deviation from $a_m(\emptyset) = FT$ to $a_m(\emptyset) = s$.■

Proof of Lemma 6: This follows using the one period payoff equations presented in Lemma 2.■

Proof of Lemma 7: $W_l(sm) > W_l(g^{FT}) > W_l(mh)$ implies $(sm) = \gamma(P_{sm}, (sm)) = \gamma(P^*, (sm))$. Moreover, $G^{EBA} \equiv \underset{S \subset N}{\cup} G(P_S, (sm)) = (sm)$ because, for $\iota \in \{s, m\}$ and $g \neq (sm)$, l deviates from $g \in \gamma(P_{\iota l}, (sm))$ to $(sm) = \gamma(P^*, (sm))$. Thus, $G(N, (sm)) = (sm)$ because l deviates from $g \neq (sm)$ to $(sm) = G(P_{sm}, (sm))$ and s and/or m cannot deviate from $(sm) \in \gamma(N, \emptyset)$ since $G^{EBA} = (sm)$. Lemma 5 showed $G(N, (sm)) \neq (sm)$ when $W_l(g^{FT}) > W_l(sm)$.■

Proof of Proposition 3: In subgames at hub–spoke networks $g = (ih)$, Lemmas 6 and 1 imply $G(N, (ih)) = g^{FT}$. In subgames at CU insider–outsider networks, $g = (ij^{CU})$, Lemmas 6 and 1 characterize $G(N, (ij^{CU}))$. Note, $G(N, (ml^{CU})) = (ml^{CU})$. In subgames at FTA insider–outsider networks $g = (ij)$, Lemmas 5 and 7 characterize $G(N, (ij))$.

Now consider the subgame at the empty network $g = \emptyset$. Using Condition 2, $\langle (ml^{CU}) \rangle \succ_m \langle g \rangle$ for $g \neq (ml^{CU})$. Using Condition 2 and $\beta \notin (\underline{\beta}_l^{Flex}, \bar{\beta}_l^{Flex})$, $\langle g^* \rangle$ is Pareto dominant for l for some $g^* \in \{(ml^{CU}), (sm)\}$. If $g^* = (ml^{CU})$ then Lemma 1 implies $G(N, \emptyset) = (ml^{CU})$ and, hence, (ml^{CU}) is the unique equilibrium path of agreements when $\langle (ml^{CU}) \rangle \succ_l \langle (sm) \rangle$.

Now let $g^* = (sm)$. Since this requires $W_l(sm) > W_l(g^{FT})$ then Lemma 7 implies $G(N, (sm)) = (sm)$. Note that $G(P_{sm}, \emptyset) = G(P_{ml}, \emptyset) = (ml^{CU})$ because i) (ml^{CU}) is Pareto dominant for m so it has a self enforcing deviation from $g \neq (ml^{CU})$ to $(ml^{CU}) = \gamma(P^*, \emptyset)$ given $\langle (ml^{CU}) \rangle \succ_l \langle g \rangle$ for $g \in \{(ml), (sl), (sl^{CU}), \emptyset\}$, ii) s has no self enforcing deviation from $(ml^{CU}) \in \gamma(P_{sm}, \emptyset)$ given $(ml^{CU}) \in \gamma(P^*, \emptyset)$ for any $a_s(\emptyset)$, and iii) l has no self enforcing deviation from $(ml^{CU}) \in \gamma(P_{ml}, \emptyset)$ given $(sm^{CU}) \in \gamma(P^*, \emptyset)$ for $a_l(\emptyset) = \phi$. Moreover, $G(P_{sl}, \emptyset) = \{(ml^{CU}), (sm)\}$ because i) $(ml^{CU}) \in \gamma(P^*, \emptyset)$ implies s has no self enforcing deviation, ii) i) $(sm^{CU}) \in \gamma(P^*, \emptyset)$ implies l has no self enforcing deviation from (ml^{CU}) and iii) l has a self enforcing deviation from $g \notin \{(ml^{CU}), (sm)\}$ to $\gamma(P^*, \emptyset) = (ml^{CU})$.

Thus, $G(N, \emptyset)$ is empty because i) m and l have a self enforcing deviation from any $g \notin \{(ml^{CU}), (sm)\}$ to $(ml^{CU}) = G(P_{ml}, \emptyset)$, ii) m has a self enforcing deviation from (sm) to $(ml^{CU}) \in G(P_{sl}, \emptyset)$ and iii) s and l have a self enforcing deviation from (ml^{CU}) to $(sm) \in G(P_{sl}, \emptyset)$. Hence, the EBAs are $\underset{S \subset N}{\cup} G(P_S, \emptyset) = \{(ml^{CU}), (sm)\}$. Thus, the equilibrium paths of agreements are (ml^{CU}) and (sm) when $\langle (sm) \rangle \succ_l \langle (ml^{CU}) \rangle$.■

Proof of Proposition 4: In subgames at hub–spoke networks $g = (ih)$, Lemmas 6 and 1 imply $G(N, (ih)) = g^{FT}$. In subgames at CU insider–outsider networks, $g = (ij^{CU})$, Lemmas 6 and 1 characterize $G(N, (ij^{CU}))$. Note, $G(N, (ml^{CU})) = (ml^{CU})$. In subgames at FTA insider–outsider networks $g=(ij)$, Lemmas 5 and 7 characterize $G(N,(ij))$.

Now consider the subgame at the empty network $g=\emptyset$. Using Condition 2, $\left\langle\left(ml^{CU}\right)\right\rangle \succ_m \langle g\rangle$ for $g\neq\left(ml^{CU}\right)$. Moreover, if $G(N,(sm))=(sm)$ then $\langle(ml)\rangle \succ_m \left\langle\left(sm^{CU}\right)\right\rangle \succ_m \langle(sm)\rangle$. Using Condition 2 and $\beta\in\left(\underline{\beta}_l^{Flex},\bar{\beta}_l^{Flex}\right)$, $\langle g^*\rangle$ is Pareto dominant for l for some $g^*\in\{(ml),(sm)\}$. Moreover, $\left\langle\left(ml^{CU}\right)\right\rangle \succ_l \langle g\rangle$ for $g\notin\left\{(ml),\left(ml^{CU}\right),(sm)\right\}\equiv G_1$.

By the logic in the proof of Proposition 3, $G(P_{ml},\emptyset)=G(P_{sm},\emptyset)=\left(ml^{CU}\right)$ and $G(P_{sl},\emptyset)\subseteq G_1$. Indeed, $(ml)\in G(P_{sl},\emptyset)$ because $\left(ml^{CU}\right)\in\gamma(P^*,\emptyset)$ for any $a_s(\emptyset)$ implies s has no self enforcing deviation while $\left(sm^{CU}\right)\in\gamma(P^*,\emptyset)$ implies l has no self enforcing deviation to $a_l(\emptyset)=\phi$ and (sm). Thus, $G(N,\emptyset)\subseteq(ml)$ because i) l has a self enforcing deviation from $g\notin G_1$ to $\left(ml^{CU}\right)=G(P_{sm},\emptyset)$, ii) s and l have a self enforcing deviation from $\left(ml^{CU}\right)$ to $(ml)=G(P_{sl},\emptyset)$, iii) l has a self enforcing deviation from (sm) to $\left(ml^{CU}\right)=G(P_{sm},\emptyset)$ when $G(N,(sm))=(mh)$ and iv) m has a self enforcing deviation from (sm) to $(ml)\in G(P_{sl},\emptyset)$ when $G(N,(sm))=(sm)$.

To conclude, begin by letting $G(N,(sm))=(mh)$. Also, let g^*_{ms} denote m's preferred PTA with s (e.g. $g^*_{ms}=(sm)$ iff $\langle(sm)\rangle \succ_s \left\langle\left(sm^{CU}\right)\right\rangle$). Note that $G(P_{sl},\emptyset)\subseteq\left\{(ml),\left(ml^{CU}\right)\right\}$. Thus, given $\langle(ml)\rangle \succ_s \left(ml^{CU}\right)$, the only possible self enforcing deviation from $(ml)\in\gamma(N,\emptyset)$ is by m to $\left(ml^{CU}\right)\in G(P_{sl},\emptyset)$. But, $\left(ml^{CU}\right)\in G(P_{sl},\emptyset)$ iff l has no self enforcing deviation to $(ml)\in\gamma(P^*,\emptyset)$ which is true iff $\langle g^*_{ms}\rangle \succ_m \langle(ml)\rangle$. Thus, $G(N,\emptyset)$ is empty and the EBAs are $\underset{S\subset N}{\cup} G(P_S,\emptyset)=\left\{\left(ml^{CU}\right),(ml)\right\}$ if $\langle g^*_{ms}\rangle \succ_m \langle(ml)\rangle$. In this case, the equilibrium paths of agreements are $\left(ml^{CU}\right)$ and (ml,sl,sm). However, $\langle(ml)\rangle \succ_m \langle g^*_{ms}\rangle$ implies $G(N,\emptyset)=(ml)$ and the unique equilibrium path of agreements is (ml,sl,sm).

Now let $G(N,(sm))=(sm)$. Note, $\langle(ml)\rangle \succ_m \left\langle\left(sm^{CU}\right)\right\rangle \succ_m \langle(sm)\rangle$ implies $G(P_{sl},\emptyset)\subseteq\{(ml),(sm)\}$ and so the only possible self enforcing deviation from $(ml)\in\gamma(N,\emptyset)$ is by s and l to $G(P_{sl},\emptyset)=(sm)$. This requires $\langle(sm)\rangle \succ_{sl} \langle(ml)\rangle$. But, $\beta\in\left(\underline{\beta}_l^{Flex},\bar{\beta}_l^{Flex}\right)$ and $\langle(sm)\rangle \succ_l \langle(ml)\rangle$ imply $\langle(ml)\rangle \succ_s \langle(sm)\rangle$. Thus, $G(N,\emptyset)=(ml)$ and the unique equilibrium path of agreements is (ml,sl,sm).■

Proof of Proposition 5: In subgames at hub–spoke networks $g=(ih)$, Lemmas 8 and 1 imply $G(N,(ih))=g^{FT}$. In subgames at CU insider–outsider networks, $g=\left(ij^{CU}\right)$, Lemmas 8 and 1 characterize $G\left(N,\left(ij^{CU}\right)\right)$. Note, $G\left(N,\left(ml^{CU}\right)\right)=\left(ml^{CU}\right)$. Lemmas 5 and 7 characterize $G(N,(sm))$ in the subgame at $g=(sm)$.

Now consider the subgame at the empty network $g=\emptyset$. Note that $\left\langle\left(ml^{CU}\right)\right\rangle \succ_m \langle g\rangle$ for any $g\neq\left(ml^{CU}\right)$ and $\left\langle\left(ml^{CU}\right)\right\rangle \succ_l \langle g\rangle$ for $g\notin\left\{\left(ml^{CU}\right),(sm)\right\}$. Moreover, $\langle(sm)\rangle \succ_l \left\langle\left(ml^{CU}\right)\right\rangle$ requires $G(N,(sm))=(sm)$.

Thus, Lemma 1 implies $G(N,\emptyset)=\left(ml^{CU}\right)$, and thus $\left(ml^{CU}\right)$ is the unique equilibrium path of agreements, if $\left\langle\left(ml^{CU}\right)\right\rangle \succ_l \langle g\rangle$ for any $g\neq\left(ml^{CU}\right)$.

Now let $\langle (sm) \rangle \succ_l \langle (ml^{CU}) \rangle$. Then $G(P_S, \emptyset) \subseteq \{(ml^{CU}), (sm)\} \equiv \tilde{G}$ for any $S \subset N$ because, for any $g \notin \tilde{G}$, m or l deviate from $g \in \gamma(P_S, \emptyset)$ to $(ml^{CU}) = \gamma(P^*, \emptyset)$. Moreover, $(ml^{CU}) \in G(P_S, \emptyset)$ for any $S \subset N$ since i) $\langle (ml^{CU}) \rangle \succ_m \langle g \rangle$ for any g, and ii) $(sm^{CU}) \in \gamma(P^*, \emptyset)$ and $(ml^{CU}) \in \gamma(P^*, \emptyset)$, respectively, deter any deviation by l or s from $(ml^{CU}) \in \gamma(P_S, \emptyset)$. Additionally, $(sm) \in G(P_{sl}, \emptyset)$ because $\langle (sm) \rangle \succ_l \langle g \rangle$ for $g \neq (sm)$ and $(ml^{CU}) \in \gamma(P^*, \emptyset)$ deters any deviation by s.

Thus, $G(N, \emptyset)$ is empty because i) m and l deviate from any $g \notin \tilde{G}$ to $(ml^{CU}) \in G(P_{ml}, \emptyset)$, ii) m deviates from (sm) to $(ml^{CU}) \in G(P_{sl}, \emptyset)$ and iii) s and l deviate from (ml^{CU}) to $(sm) \in G(P_{sl}, \emptyset)$. Hence, the EBAs are $\underset{S \subset N}{\cup} G(P_S, \emptyset) = \{(ml^{CU}), (sm)\}$. Thus, the equilibrium paths of agreements are (ml^{CU}) and (sm) when $\langle (sm) \rangle \succ_l \langle (ml^{CU}) \rangle$.■

Proof of Proposition 7: Given $W_l(g^{FT}) > W_l(\emptyset)$ and the construction of $\bar{\tau}^i((ij^M))$, $\gamma(P^*, \emptyset) = \{\emptyset, g^{FT}, (sm^M), (sl^M), (ml^M)\}$ and $W_i(g^{FT}) - W_i(ij^M) \propto -1.4443\alpha_i^2 + 1.111\alpha_j^2 + 1.68\alpha_k^2 > 0$ for any $i, j. W_l(g^{FT}) > W_l(\emptyset)$. Also, $W_i(g^{FT}) - W_i(jk^M) \propto -1.8\alpha_i^2 + 1.111(\alpha_j^2 + \alpha_k^2)$ implies $W_s(g^{FT}) > W_s(ml^M)$.

Two observations establish the proof. First, $g^{FT} \in G(P_{sm}, \emptyset)$ and, by similar logic, $g^{FT} \in G(P_{sl}, \emptyset)$. $g^{FT} \in G(P_{sm}, \emptyset)$ follows because i) $g^{FT} \in \gamma(P_{sm}, \emptyset)$ and ii) $\emptyset \in \gamma(P^*, \emptyset)$ deters m's deviation to $a_m(\emptyset) = \phi$ and (sl^M). Second, $(sm^M) \notin G(P_{sm}, \emptyset)$ and, by similar logic, $(sl^M) \notin G(P_{sl}, \emptyset)$. To see $(sm^M) \notin G(P_{sm}, \emptyset)$, note that m has a self enforcing deviation from $a_m(\emptyset) = s^M$ to $a_m(\emptyset) = l^M$ and $\gamma(P^*, \emptyset) = (ml^M)$ iff $\langle (ml^M) \rangle \succ_m \langle (sm^M) \rangle$ which holds unless $G(N, (sm^M)) = g^{FT}$ and $G(N, (ml^M)) = (ml^M)$. However, given $W_l(g^{FT}) > W_l(ml^M)$ and $W_m(g^{FT}) > W_m(ml^M)$, $G(N, (ml^M)) = (ml^M)$ requires $W_s(ml^M) > W_s(g^{FT})$ which is false.

Given $\langle g^{FT} \rangle \succ_i \langle (ij^M) \rangle \succ_i \langle \emptyset \rangle$ for all i, j and $\langle g^{FT} \rangle \succ_s \langle (ml^M) \rangle$, the only possible deviations from $g^{FT} \in \gamma(N, \emptyset)$ are m to (sl^M) and l to (sm^M). But these deviations are not self enforcing since $(sl^M) \notin G(P_{sl}, \emptyset)$ and $(sm^M) \notin G(P_{sm}, \emptyset)$. Thus, $G(N, \emptyset) \supseteq g^{FT}$. Indeed, $G(N, \emptyset) = g^{FT}$ because of self enforcing deviations by s and m from $g \in \{\emptyset, (sm^M), (ml^M)\}$ to $g^{FT} \in G(P_{sm}, \emptyset)$ and s and l from (sl^M) to $g^{FT} \in G(P_{sl}, \emptyset)$.■

C Full equilibrium characterization under a small degree of asymmetry

Notation wise, let g_{ij}^* denote i's preferred PTA from $\{(ij), (ij^{CU})\}$. In characterizing the equilibrium, the following conditions are useful. When Condition 4 fails then, even though m prefers an FTA with l over any PTA with s, l will have no self enforcing deviation from $a_l(\emptyset) = FT$ to $a_l(\emptyset) = m$, i.e. $(ml) \notin G(P_{sm}, \emptyset)$, because m has a subsequent self enforcing deviation from $a_m(\varnothing) = l$ to $a_m(\varnothing) = l^{CU}$.

Condition 4. i) $\langle g^*_{ms}\rangle \succ_m \langle (ml)\rangle$ if $\langle (ml)\rangle \succ_l \langle g^{FT}\rangle$ or ii) $\langle (ml^{CU})\rangle \succ_m \langle g^*_{ms}\rangle$ and $\langle (sl)\rangle \succ_l \langle (ml^{CU})\rangle$

Condition 5 provides the necessary and sufficient conditions under which l has no self enforcing deviation from $a_l(\varnothing) = FT$ to $a_l(\varnothing) = m$ when Condition 4 holds.

Condition 5. i) $\langle (sl)\rangle \succ_l \langle (ml^{CU})\rangle$, ii) $\langle (ml)\rangle \succ_s \langle (sm)\rangle$ or $\langle (ml)\rangle \succ_m \langle (sm)\rangle$, and iii) $\langle (ml)\rangle \succ_m \langle (sm^{CU})\rangle$.

In characterizing the equilibrium path of agreements, let $\Theta_1^{G^{EBA}} \equiv \{g^{FT}, (ml, sl, sm)\}$ and
$\Theta_2^{G^{EBA}} \equiv \{g^{FT}, (ml, sl, sm), (sl, ml, sm), (ml^{CU}, g^{FT}), (sl^{CU}, g^{FT})\}$. The superscript G^{EBA} denotes that these are the network paths associated with the EBAs $G^{EBA} \equiv \underset{S\subset N}{\cup} G(P_S, \varnothing)$. Following Proposition 8 is a sketch of the proof.

Proposition 8. *Assume Condition 1 holds. If either i)* $\langle g^{FT}\rangle \succ_l \langle (ml)\rangle$*, ii) Condition 4 fails, or iii) Condition 4 holds but Condition 5 fails, the unique equilibrium path of agreements is* g^{FT}*. Otherwise, the set of equilibrium paths of agreements is* $\Theta_1^{G^{EBA}}$ *if* $\langle (ml)\rangle \succ_m \langle g^*_{ms}\rangle$ *and* $\Theta_2^{G^{EBA}}$ *otherwise.*

Proof. In subgames at hub–spoke networks $g = (ih)$, Lemmas 4 and 1 imply $G(N, (ih)) = g^{FT}$. In subgames at CU insider–outsider networks, $g = (ij^{CU})$, Lemmas 4 and 1 imply $G(N, (ij^{CU})) = g^{FT}$. In subgames at FTA insider–outsider networks $g = (ij)$, Lemma 5 characterizes $G(N, (ij))$.

Now consider the subgame at the empty network $g = \varnothing$. Begin by considering whether or not $G(N, \varnothing)$ is empty. $g \notin G(N, \varnothing)$ for $g \in \{\varnothing, (sl^{CU}), (sm^{CU}), (ml^{CU})\}$ because $\langle g^{FT}\rangle \succ_N \langle g\rangle$ for $g \neq g^{FT}$ implies $g \notin \gamma(N, \varnothing)$. Moreover, $g \notin G(N, \varnothing)$ for $g \in \{(ml), (sl)\}$ because $\langle g^{FT}\rangle \succ_{sm} \langle g\rangle$ implies s and m have a self enforcing deviation from g to $g^{FT} \in G(P_{sm}, \varnothing)$. $g^{FT} \in G(P_{sm}, \varnothing)$ follows because i) $\langle g^{FT}\rangle \succ_s \langle g\rangle$ for $g \neq g^{FT}$, ii) $\langle g^{FT}\rangle \succ_m \langle g\rangle$ for $g \notin \{g^{FT}, (sm)\}$, and iii) $(sl^{CU}) \in \gamma(P^*, \varnothing)$ deters m's deviation to (sm). Hence, $G(N, \varnothing) \subseteq \{g^{FT}, (sm)\}$.

To see that $(sm) \notin G(N, \varnothing)$, and thus $G(N, \varnothing) \subseteq g^{FT}$, first let $G(N, (sm)) = g^{FT}$. Then s and m deviate from (sm) to $g^{FT} \in G(P_{sm}, \varnothing)$. Second, let $G(N, (sm)) = (mh)$ noting that $\langle g^{FT}\rangle \succ_l \langle g^*_{ms}\rangle$. Suppose $\langle g^{FT}\rangle \succ_l \langle (ml)\rangle$. Then s and l deviate from (sm) to $g^{FT} \in G(P_{sl}, \varnothing)$ given $\langle g^{FT}\rangle \succ_{sl} \langle g\rangle$ for $g \neq g^{FT}$ and $W_m(g^{FT}) > W_m(\varnothing)$. Now suppose $\langle (ml)\rangle \succ_l \langle g^{FT}\rangle$. If $\langle g^*_{lm}\rangle \succ_m \langle g^*_{ms}\rangle$, then $\langle g^{FT}\rangle \succ_m \langle (ml)\rangle \succ_m \langle (sm)\rangle$ which implies $(sm) \notin G(N, \varnothing)$ since s and m deviate from (sm) to $g^{FT} \in G(P_{sm}, \varnothing)$. If $\langle g^*_{ms}\rangle \succ_m \langle (ml)\rangle$, then s and l deviate from (sm) to $g^{FT} \in G(P_{sl}, \varnothing)$.

Therefore, $G(N, \emptyset) \subseteq g^{FT}$ and the EBAs are $G(N, \emptyset) = g^{FT}$ or $G^{EBA} \equiv \underset{S \subset N}{\cup} G(P_S, \emptyset)$ depending on whether there is a self enforcing deviation from g^{FT}. By part ii) of Lemma 2, there is no profitable joint deviation. Given $\langle g^{FT} \rangle \succ_s \langle g \rangle$ for $g \neq g^{FT}$, consider unilateral deviations by m and l. m has no unilateral deviation since $\langle (sm) \rangle \succ_m \langle g^{FT} \rangle$ implies $\langle (sl) \rangle \succ_l \langle (sm) \rangle$ and thus, given $\langle (sl) \rangle \succ_s \langle (sm) \rangle$, $(sm) \notin \gamma(P_{sl}, \emptyset)$. Noting that i) $(sl) \notin G(P_{sm}, \emptyset)$ because m deviates to $(ml) \in \gamma(P^*, \emptyset)$ and ii) $\langle g^{FT} \rangle \succ_l \langle g \rangle$ for $g \notin \{g^{FT}, (ml), (sl)\}$, then only unilateral deviations by l warrant consideration.

Conditions 4 and 5 determine whether the EBAs are $G(N, \emptyset) = g^{FT}$ or $G^{EBA} \equiv \underset{S \subset N}{\cup} G(P_S, \emptyset)$ and, in the latter case, also determine G^{EBA}.■

References

Aghion, P., Antràs, P., Helpman, E., 2007. Negotiating free trade. Journal of International Economics 73 (1), 1–30.

Bagwell, K., Staiger, R., 1997a. Multilateral tariff cooperation during the formation of customs unions. Journal of International Economics 42 (1), 91–123.

Bagwell, K., Staiger, R., 1997b. Multilateral tariff cooperation during the formation of free trade areas. International Economic Review 38 (2), 291–320.

Bagwell, K., Staiger, R. W., 2010. Backward stealing and forward manipulation in the WTO. Journal of International Economics 82 (1), 49–62.

Baier, S. L., Bergstrand, J. H., 2004. Economic determinants of free trade agreements. Journal of International Economics 64 (1), 29–63.

Bernheim, B., Peleg, B., Whinston, M., 1987. Coalition-proof Nash equilibria I. concepts. Journal of Economic Theory 42 (1), 1–12.

Bhagwati, J., 1991. The world trading system at risk. Princeton University Press.

Bhagwati, J., 1993. Regionalism and multilateralism: an overview. In: Panagriya, A., de Melo, J. (Eds.), New dimensions in regional integration. World Bank.

Bond, E. W., Riezman, R. G., Syropoulos, C., 2004. A strategic and welfare theoretic analysis of free trade areas. Journal of International Economics 64 (1), 1–27.

Chen, M., Joshi, S., 2010. Third-country effects on the formation of free trade agreements. Journal of International Economics 82 (2), 238–248.

Diamantoudi, E., 2003. Equilibrium binding agreements under diverse behavioral assumptions. Economic Theory 22 (2), 431–446.

Dutta, B., Ghosal, S., Ray, D., 2005. Farsighted network formation. Journal of Economic Theory 122 (2), 143–164.

Eichengreen, B., Irwin, D., 1998. The role of history in trade flows. In: M. Canzoneri, W. J. E., Grilli, V. (Eds.), The Regionalization of the World Economy. University of Chicago Press.

Estevadeordal, A., Freund, C., Ornelas, E., 2008. Does regionalism affect trade liberalization toward nonmembers? The Quarterly Journal of Economics 123 (4), 1531–1575.

Facchini, G., Silva, P., Willmann, G., 2012. The customs union issue: Why do we observe so few of them? Journal of International Economics.

Freund, C., McLaren, J., 1999. On the dynamics of trade diversion: Evidence from four trade blocks. Board of Governors of the Federal Reserve System, International Finance Discussion Paper No. 637.

Furusawa, T., Konishi, H., 2007. Free trade networks. Journal of International Economics 72 (2), 310–335.

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

Grossman, G., Helpman, E., 1995. The politics of free-trade agreements. The American Economic Review 85.

Jackson, M., 2008. Social and economic networks. Princeton University Press.

Kennan, J., Riezman, R., 1990. Optimal tariff equilibria with customs unions. Canadian Journal of Economics, 70–83.

Krishna, P., 1998. Regionalism and multilateralism: A political economy approach. Quarterly Journal of Economics 113 (1), 227–251.

Krugman, P., 1991. The move toward free trade zones. policy implications of trade and currency zones: A symposium sponsored by the federal reserve bank of kansas city. Federal Reserve Bank of Kansas City, Kansas City, 7–41.

Lake, J., 2013. Free trade agreements as dynamic farsighted networks. Departmental Working Papers 1304, Southern Methodist University, Department of Economics.

Ludema, R., 1991. International trade bargaining and the Most-Favored-Nation clause. Economics & Politics 3 (1), 1–20.

McLaren, J., 2002. A theory of insidious regionalism. Quarterly Journal of Economics 117 (2), 571–608.

Melatos, M., Dunn, S., 2013. Flexibility in trade bloc design. Mimeo.

Melatos, M., Woodland, A., 2007. Endogenous trade bloc formation in an asymmetric world. European Economic Review 51 (4), 901–924.

Missios, P., Saggi, K., Yildiz, H., 2013. The implications of alternative forms of bilateralism for multilateralism: Free trade agreements vs. customs unions. Mimeo.

Mrázová, M., Vines, D., Zissimos, B., 2012. Is the GATT/WTO's article XXIV bad? Journal of International Economics.

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

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

Ornelas, E., 2007. Exchanging market access at the outsiders' expense: the case of customs unions. Canadian Journal of Economics 40 (1), 207–224.

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

Ornelas, E., Liu, X., 2012. Free trade agreements and the consolidation of democracy. Mimeo.

Ray, D., Vohra, R., 1997. Equilibrium binding agreements. Journal of Economic Theory 73 (1), 30–78.

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

Roberts, M., Tybout, J., 1997. The decision to export in colombia: an empirical model of entry with sunk costs. American Economic Review 87 (4), 545–564.

Saggi, K., Woodland, A., Yildiz, H. M., 2013. On the relationship between preferential and multilateral trade liberalization: the case of customs unions. American Economic Journal: Microeconomics 5 (1), 63–99.

Saggi, K., Yildiz, H., 2010. Bilateralism, multilateralism, and the quest for global free trade. Journal of International Economics 81 (1), 26–37.

Saggi, K., Yildiz, H., 2011. Bilateral trade agreements and the feasibility of multilateral free trade. Review of International Economics 19 (2), 356–373.

Seidmann, D., 2009. Preferential trading arrangements as strategic positioning. Journal of International Economics 79, 143–159.

Syropoulos, C., 2002. On tariff preferences and delegation decisions in Customs Unions: A Heckscher–Ohlin approach. The Economic Journal 112 (481), 625–648.

Syropoulos, C., 2003. Rules for the disposition of tariff revenues and the determination of common external tariffs in Customs Unions. Journal of International Economics 60 (2), 387–416.

WTO, 2011. World Trade Report 2011. The WTO and preferential trade agreements: From co-existence to coherence.