Free Trade Agreements as dynamic farsighted networks

James Lake*
Southern Methodist University

May 2015

Abstract

In the presence of multilateral negotiations, are Free Trade Agreements (FTAs) necessary for, or will they prevent, global free trade? I answer this question using a dynamic farsighted model of network formation among asymmetric countries. Ultimately, FTAs prevent global free trade when there are two larger countries and one smaller country but FTAs can be necessary for global free trade when there are two smaller countries and one larger country. The model provides insights into the dynamics of recent trade negotiations involving the US and recent results in the literature on the empirical determinants of trade agreements.

JEL: C71, F12, F13

Keywords: Free Trade Agreements, FTA exclusion incentive, preference erosion, multilateralism, global free trade, networks, farsighted

---

*Department of Economics, Southern Methodist University, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas, TX 75275, email: jlake@smu.edu. I would like to thank Pravin Krishna, M. Ali Khan, Kamal Saggi, Emanuel Ornelas, Sumit Joshi, Han Ozsoylev, Saltuk Ozerturk, Murat Yildiz, Ray Riezman, Sanjeev Goyal, Alan Woodland and Martin Richardson for useful comments and discussion as well as seminar and conference participants at many seminars and conferences.

1 Introduction

Recent decades have witnessed an unprecedented proliferation of Free Trade Agreements (FTAs). Although sanctioned by the WTO in GATT Article XXIV, FTAs are discriminatory by construction and contradict the central principle of non–discrimination articulated in the Most Favored Nation (MFN) principle of GATT Article I.[1] Thus, the proliferation of FTAs has stimulated substantial debate on whether FTAs hinder or facilitate greater liberalization, especially given the lack of multilateral liberalization since the 1994 Uruguay Round. That is, are FTAs "building blocs" or "stumbling blocs" to global free trade?

In essence, this is a dynamic issue concerning the evolution of trade agreements over time yet much of the literature uses static three country models. Many papers ask if an arbitrarily chosen pair of countries want to form an FTA and, if so, how this affects incentives for expansion of the agreement to include the third country, thus achieving global free trade (e.g., Levy (1997), Krishna (1998), Ornelas (2005a,b)). However, Saggi and Yildiz (2010, 2011) emphasize the importance of comparing the equilibrium outcomes of two games. They compare a "bilateralism game", where countries choose between forming bilateral FTAs or moving directly to global free trade, to a "multilateralism game", where countries cannot form FTAs.[2] I follow this approach in a three country dynamic network formation framework and, by comparing the equilibrium of these different games, classify FTAs as "strong building blocs" if global free trade is only attained in the presence of FTAs but "strong stumbling blocs" if global free trade is only attained in the absence of FTAs.

Viewing links between players as trade agreements between countries, the dynamic network formation model has three defining features. First, at most one agreement can form in a period. That is, I interpret a period as the length of time needed to complete FTA negotiations; in practice, completion of FTA negotiations typically takes many years.[3] Second, agreements formed in previous periods are binding. Ornelas (2008, p.218) and Ornelas and Liu (2012, p.13), among others, have argued the binding nature of trade agreements is pervasive in the literature and realistic.[4] Third, in the spirit of Aghion et al. (2007), I impose a protocol where, in each period, a "leader" country can make trade agreement proposals to the "follower" countries. However, unlike Aghion et al. (2007), I allow the follower coun-

[1] GATT Article I requires any tariff reductions afforded to one country are afforded to all countries. But GATT Article XXIV provides an escape clause: FTA members can eliminate tariffs between themselves if they do not raise tariffs or non tariff barriers on other countries.

[2] This approach was first adopted by Riezman (1999).

[3] For example, NAFTA diplomatic negotiations date back to 1988 (Odell (2006, p.193)) despite the agreement being implemented in 1994.

[4] They argue realism both from the perspective of practical observation and as a reduced form for a more structural explanation. For example, see McLaren (2002) for sunk costs as an explanation and, among others, Freund and McLaren (1999) for empirical support.

tries to make proposals in periods where the proposal of the leader country is rejected or the leader chooses to make no proposal. Within this dynamic network formation framework, countries are farsighted because they base their actions on the continuation payoff of forming an agreement rather than the one period payoff.

To solve the equilibrium path of network formation, I posit a general specification of one period payoffs rather than assume a particular trade model. The essence of the specification, although the exact conditions are weaker, is twofold. First, FTAs benefit members but may harm non–members. Second, and most importantly, a pair of "insider" countries (i.e. countries who have the sole FTA in existence) hold an "FTA exclusion incentive": in terms of their one period payoff, insiders want to exclude the "outsider" country from a direct move to global free trade. Although not present in all standard trade models, I show FTA exclusion incentives arise in numerous models. Moreover, Section 6 discusses how an observable implication of FTA exclusion incentives finds empirical support in Chen and Joshi (2010).

The FTA exclusion incentive creates an important dynamic tradeoff for an insider. Further FTA formation allows it to become the "hub" and benefit from the additional reciprocal preferential access exchanged with the outsider. However, the would–be hub anticipates a subsequent FTA between the "spoke" countries which erodes the value of reciprocal preferential access enjoyed as the hub. That is, fears of *preference erosion* undermine an insider's incentive to engage in subsequent FTA formation. The FTA exclusion incentive, and the underlying fear of preference erosion, fundamentally affects the role played by FTAs.[5]

When countries are symmetric, FTAs can be strong stumbling blocs but not strong building blocs. On one hand, global free trade is attained in the multilateralism game because each country views the world market as attractive enough that it does not veto a direct move to global free trade. On the other hand, global free trade is attained in the bilateralism game only when the discount factor is sufficiently small because then an insider's fear of future preference erosion is small relative to the myopic attractiveness of exchanging further reciprocal market access and becoming the hub. Thus, FTAs are strong stumbling blocs and prevent global free trade when the discount factor exceeds a threshold.

More generally, the role of FTAs depends crucially on asymmetries. To model asymmetry, I assume countries can be ranked in terms of their "attractiveness" as FTA partners based on a scalar parameter α. This interpretation includes, among others, market size or technology asymmetries. For want of better terminology, I interpret countries with a higher α as being "larger". FTAs are strong stumbling blocs, preventing global free trade, with two larger countries and one smaller country. Here, the largest country views the world market as

[5]Hoekman (2006) and Limão and Olarreaga (2006) investigate empirical and policy implications of preference erosion.

attractive enough that it does not veto global free trade in the multilateralism game. But, the value of reciprocal preferential access protected by the two largest countries as insiders is high which generates a strong fear of preference erosion (i.e. low critical discount factor) and prevents global free trade from emerging in the bilateralism game for a large range of the discount factor. Thus, FTAs are strong stumbling blocs and prevent global free trade with two (sufficiently) larger countries and one (sufficiently) smaller country.

In contrast, FTAs are strong building blocs and necessary for global free trade with two smaller countries and one larger country. Here, the largest country views the world market as relatively unattractive and vetoes global free trade in the multilateralism game. However, as insiders, the value of reciprocal preferential access protected by the two largest countries (i.e. the largest country and the biggest smaller country) is low. As such, preference erosion fears are weak (i.e. high critical discount factor). Thus, global free trade emerges in the bilateralism game over a large range of the discount factor. Hence, FTAs are strong building blocs and necessary for global free trade with two (sufficiently) smaller countries and one (sufficiently) larger country.

Importantly, the model helps shed some light on real world FTA formation and non-formation. The model relates the path of FTA formation to (i) country asymmetries (matching empirical evidence of Chen and Joshi (2010)) and (ii) the order FTA negotiations commence. These predictions are consistent with recent US negotiations involving numerous partners. Moreover, as discussed above, an observable implication stemming from the result that FTA exclusion incentives, and the underlying fear of preference erosion, drive FTA non-formation receives empirical support from Chen and Joshi (2010).

While they do not refer to it as an FTA exclusion incentive or a fear of preference erosion, Mukunoki and Tachi (2006) identify the associated trade off faced by FTA insiders. But, they do not address the strong building bloc–strong stumbling bloc issue nor do they model country asymmetries. Indeed, Krugman (1991), Grossman and Helpman (1995) and Saggi and Yildiz (2010, p.27) have emphasized the importance of country asymmetries. To this end, my model delivers a clear and intuitive explanation linking country asymmetries and the role of FTAs as strong building blocs or strong stumbling blocs.

The strong stumbling bloc role of FTAs is the most important difference with Saggi and Yildiz (2010, 2011). Not only is their static framework unable to capture the dynamic farsighted logic of preference erosion, but their underlying trade model does not exhibit FTA exclusion incentives which are crucial to my strong stumbling bloc result.[6] While Saggi et al.

[6]Saggi and Yildiz (2010) use the popular "competing exporters model" with endowment asymmetry. Interestingly, this setting does not feature FTA exclusion incentives but Lemma 1 here will show that the competing exporters model with market size asymmetry does feature FTA exclusion incentives.

(2013) find that Customs Unions (CUs) can be strong stumbling blocs, the WTO requirement that CU members impose a common tariff on non–members implies CUs are very different types of agreements than FTAs. Moreover, FTAs make up 90% of all preferential trade agreements (i.e. FTAs and CUs) which places utmost importance on the FTA analysis.[7]

Using network formation models to address FTA formation dates back to Goyal and Joshi (2006). In a symmetric oligopolistic setting they show the complete network (i.e. global free trade) is pairwise stable (Jackson and Wolinsky (1996)) and the unique efficient network. Furusawa and Konishi (2007) employ a model with a continuum of differentiated goods and show that, when consumers view goods as unsubstitutable, the pairwise stable network involves an FTA between two countries if and only if the countries have a similar level of industrialization (i.e. similar number of firms). Using a dynamic, but myopic best response, network formation model, Zhang et al. (2014) show the attainment of global free trade can hinge on the special case of three countries.

Finally, my model shares similar features to the three country dynamic model of Seidmann (2009), but the question of interest differs. His interest lies in whether the equilibrium type of trade agreement is a CU or an FTA. But, my interest rests on whether global free trade is eventually attained which is a moot issue for Seidmann (2009) because transfers imply global free trade always emerges in equilibrium since it is efficient (i.e. maximizes world welfare). In contrast, I assume transfers are not available to countries so global free trade need not obtain even if global free trade is efficient.[8]

2 Payoffs

In this section, I devote significant effort to develop general properties on one period and continuation payoffs that fit a variety of underlying trade models but are also sufficient to explicitly solve the equilibrium path of networks. Nevertheless, to fix the basic intuition of these general properties, I initially discuss an extremely simple underlying trade model: the oligopolistic model with a common exogenous tariff and a government objective function that only depends on firm profits (hereafter, the "political economy oligopolistic model"). However, Section 2.2 shows the general properties fit numerous trade models and, thus, the important intuition emerges in numerous trade models.

Before proceeding, some notation and terminology is needed. The set of countries is

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

[8] According to Bagwell and Staiger (2010, p.50), reality is "... positioned somewhere in between the extremes of negotiations over tariffs only and negotiations over tariffs and [transfers]...". Aghion et al. (2007) and Bagwell and Staiger (2010) allow transfers while others including Riezman (1999), Furusawa and Konishi (2007), and Saggi and Yildiz (2010) do not.

$N = \{s, m, l\}$ and g denotes a network of trade agreements. Figure 1 illustrates the possible networks and terminology. Generally, a link between two nodes indicates an FTA. But, the free trade network could represent either three FTAs or a three country MFN agreement.

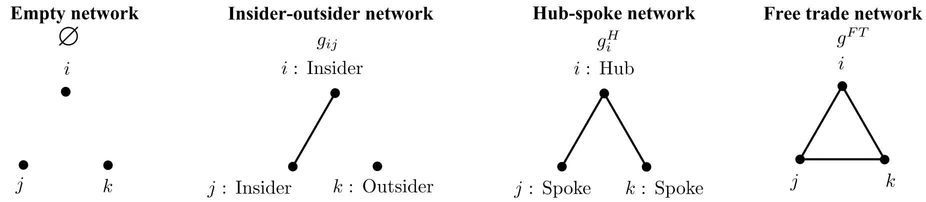

Figure 1: Networks and position terminology

2.1 A simple model and general properties

To fix ideas, consider an oligopolistic model. Three countries, each with a single firm, produce a homogenous good in segmented international markets. x_{ij} denotes the quantity sold by country i in country j's market (this allows $j = i$). Country i's demand is $d_i(p_i) = \bar{d}_i - p_i$ where $\bar{d}_i$ denotes country i's market size and p_i denotes the price in country i. Ruling out prohibitive tariffs, country i imposes a tariff τ_{ij} on country j (naturally, $\tau_{ii} = 0$).

Assuming a common and constant marginal cost (normalized to zero), country i's maximization problem in country j has the standard form: $\max\limits_{x_{ij}} \left[\left(\bar{d}_j - \sum_{j \in N} x_{ij}\right) - \tau_{ji}\right] x_{ij}$. Given a network g, the equilibrium quantity $x^*_{ij}(g)$ is

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

where (i) $\eta_j(g)$ is the number of countries facing a zero tariff in country j (including country j itself) and, per WTO rules, (ii) $\bar{\tau}_j(g)$ is the non-discriminatory tariff faced by countries who do not have an FTA with country j, and (iii) $\tau_{ji}(g) = 0$ if i and j have an FTA. Country i's equilibrium profits in country j are $\pi_{ij}(g) = \left(x^*_{ij}(g)\right)^2$ and country i's total profits are $\pi_i(g) = \sum_{j \in N} \pi_{ij}(g)$.

Defining government i's one period payoff from a network g as $v_i(g) = \pi_i(g)$, four properties succinctly summarize the payoff structure of the symmetric political economy oligopolistic model. First, the reciprocal exchange of preferential market access makes FTAs mutually beneficial for members: $v_h(g + ij) > v_h(g)$ for $h = i, j$ where $g + ij$ denotes the network that adds the FTA between countries i and j to g. Second, the reciprocal exchange of preferential access by FTA members hurts non-members: $v_k(g) > v_k(g + ij)$ for $k \neq i, j$. Third, when moving from no agreements to global free trade, the gains from exchanging pairwise preferential access with the other two countries outweighs the preferential access exchanged by the other countries between themselves so that global free trade benefits each country: $v_i\left(g^{FT}\right) > v_i(\emptyset)$ for all i. However, fourth, the additional gains for an insider when forming an additional FTA and becoming the hub are dominated by the losses suffered upon a subsequent spoke-spoke FTA: $v_i\left(g_i^H\right) > v_i(g_{ij}) > v_i\left(g^{FT}\right)$. Importantly, $v_i(g_{ij}) > v_i\left(g^{FT}\right)$ represents the "FTA exclusion incentive" whereby insiders want to exclude the outsider from expansion to global free trade to avoid eroding their reciprocal preferential market access.

2.1.1 One period payoffs under symmetry

The assumptions in the political economy oligopolistic model that governments only care about firm profits and set tariffs exogenously are problematic. In particular, the model cannot generate "tariff complementarity" whereby FTA formation induces members to lower tariffs on non–members.[9] Indeed, in many trade models, tariff complementarity is strong enough that FTA formation benefits the non-member despite the inherent discrimination. Nevertheless, later results do not rely on the presence or absence of tariff complementarity. Under symmetry, later results only rely on the following one period payoff properties.

Condition 1. Countries are symmetric and
(i) $v_h(g+ij) > v_h(g)$ for $h=i,j$ and $v_h\left(g^{FT}\right) > v_h(\emptyset)$ for any h
(ii) $v_h\left(g_i^H\right) > v_h\left(g^{FT}\right)$ for any h
(iii) $v_h(g_{ij}) > v_h\left(g^{FT}\right)$ for any h

While parts (i)-(iii) were described above in Section 2, part (ii) has more generality: FTA formation need only impose negative externalities on the non-member if the FTA is a spoke-spoke FTA. That is, part (ii) allows the non-member to benefit from tariff complementarity.[10,11] Thus, the important properties driving later results under symmetry do not depend on the primary limitations of the political economy oligopolistic model.

2.1.2 One period payoffs under asymmetry

To begin, it is useful to discuss how market size asymmetry affects Condition 1 in the political economy oligopolistic model. The basic intuition is twofold.

[9] Theoretically, see, for example, Ornelas (2005a) and Saggi and Yildiz (2010). Empirically, see Estevadeordal et al. (2008).

[10] Note, by construction, there is no tariff complementarity upon a spoke-spoke FTA because the spokes already set zero tariffs on the hub.

[11] Note, part (iii) has a subtle implication for the degree of tariff complementarity upon FTA formation at the insider–outsider network: the non-member, i.e. the insider–turned–spoke, cannot benefit from tariff complementarity. To see this, note that $v_i\left(g_j^H\right) > v_i(g_{ij})$ together with part (i) would imply $v_i\left(g^{FT}\right) > v_i\left(g_j^H\right) > v_i(g_{ij})$. Nevertheless, this implication is rather weak. Unlike when FTA formation occurs at the empty network and the non-member can benefit from tariff complementarity in *both* member markets, an insider–turned–spoke can *only* benefit from tariff complementarity in the outsider's market.

First, FTA formation is more attractive with a larger partner due to the greater market access gained in the partner market: $v_k(g+ik) > v_k(g+jk)$ if and only if $\bar{d}_i > \bar{d}_j$. However, this implies smaller countries may not hold FTA exclusion incentives against larger countries because the significant degree of market access gained with the larger country could outweigh erosion of the preferential access shared by smaller insider countries.

Second, FTA formation is less attractive for a larger country because of the greater domestic market access conceded. Thus, a larger country may suffer from FTA formation: $v_i(g+ij) \gtrless v_i(g)$ if $\bar{d}_i > \bar{d}_j$. Further, relative to no agreements, the largest country may not benefit from global free trade: $v_i\left(g^{FT}\right) \gtrless v_i(\emptyset)$ if $\bar{d}_i = \max\left\{\bar{d}_i, \bar{d}_j, \bar{d}_k\right\}$. This twofold intuition implies Condition 1 needs weakening under asymmetry.

Under asymmetry, Condition 2 weakens the properties of Condition 1. Country asymmetry is simply modeled by letting α_i denote country i's characteristics. In the oligopolistic model above, α_i represents the intercept on country i's demand curve $\bar{d}_i$. However, later I will interpret α_i as endowment or technology characteristics of country i.[12]

Condition 2. When countries are asymmetric, Condition 1 holds except that

(i) $v_k(g+ik) > v_k(g+jk)$ if and only if $\alpha_i > \alpha_j$

(ii) $v_i\left(g_j^H\right) \gtrless v_i(g_{jk})$ if $\alpha_i > \alpha_j$

(iii) $v_i(g_{ik}) \gtrless v_i(\emptyset)$ if $\alpha_i > \alpha_j > \alpha_k$

(iv) $v_i\left(g^{FT}\right) \gtrless v_i(\emptyset)$ if $\alpha_i > \alpha_j > \alpha_k$ with $\frac{\partial\left[v_i\left(g^{FT}\right)-v_i(\emptyset)\right]}{\partial(\alpha_j/\alpha_k)} > 0$ and $\frac{\partial\left[v_i\left(g^{FT}\right)-v_i(\emptyset)\right]}{\partial(\alpha_i/\alpha_k)} < 0$

(v) $v_i(g_{ij}) \gtrless v_i\left(g^{FT}\right)$ if $\alpha_i = \min\{\alpha_i, \alpha_j, \alpha_k\}$ or $\alpha_j = \min\{\alpha_i, \alpha_j, \alpha_k\}$

(vi) $v_h\left(g^{FT}\right) > v_h(g_{jk})$ for $h = j, k$ if $\alpha_i > \alpha_j > \alpha_k$ and $v_i(g_{jk}) > v_i\left(g_j^H\right)$

Although somewhat tedious, Condition 2 follows naturally from Condition 1 in four ways.

First, part (i) states the way asymmetry is modeled.[13] Second, like discussed above, parts (ii)-(iii) weaken the extent that FTAs mutually benefit members. Part (ii) says an outsider may not benefit from FTA formation with a "less attractive" country. Part (iii) says the "most attractive" country may not benefit from becoming an insider with the "least attractive" country. Third, part (iv) says the most attractive player may prefer the status quo of no agreements over global free trade with the following tension underlying this preference: global free trade trade becomes more (less) appealing as the moderately attractive country j (most attractive country i) becomes more attractive relative to the least attractive country k. The first part of this trade–off has the spirit of part (i). Intuitively, the second part says the characteristic that makes a country an attractive partner actually makes global free trade less appealing to the most attractive country.

[12]The α_i's could be vectors of characteristics if a mapping reduces the vector to a scalar summary statistic.

[13]Note, in common underlying trade models (e.g. oligopoly, competing exporter or competing importer models), the perception of what makes another country attractive is independent of a country's perspective.

Fourth, as discussed above, parts (v) and (vi) weaken the extent that insiders hold FTA exclusion incentives. Part (v) says insiders may not hold FTA exclusion incentives when the least attractive country is an insider.[14] Thus, FTA exclusion incentives need only exist when the two most attractive countries are insiders. Part (vi) says the insiders j and k do not hold FTA exclusion incentives when i prefers remaining the outsider over subsequent FTA formation with the most attractive insider. Given part (ii) says this is only possible when the outsider is the most attractive country, part (vi) is quite weak: if the most attractive country i is attractive enough that it does not benefit from FTA formation with the most attractive insider j, then insiders do not want to exclude the most attractive country i.

2.1.3 Continuation payoffs

Given Conditions 1 and 2 impose relatively weak conditions on one period payoffs, Condition 3 introduces properties on continuation payoffs to help characterize the equilibrium. δ denotes the discount factor.

Condition 3. (i) $v_i(g_{ij}) + \delta v_i\left(g_j^H\right) + \frac{\delta^2}{1-\delta} v_i\left(g^{FT}\right) > \frac{1}{1-\delta} v_i(\varnothing)$ if $\alpha_i \leq \alpha_j$

(ii) $v_i(g_{ij}) + \delta v_i\left(g_i^H\right) + \frac{\delta^2}{1-\delta} v_i\left(g^{FT}\right) > \frac{1}{1-\delta} v_i(\varnothing)$ if $\alpha_i \geq \alpha_j \geq \alpha_k$

(iii) $v_h(g_{ih}) + \delta v_h\left(g_i^H\right) + \frac{\delta^2}{1-\delta} v_h\left(g^{FT}\right) > \frac{1}{1-\delta} v_h(g_{jk})$ for $h = j, k$ if $\alpha_i > \alpha_j > \alpha_k$, $v_i(g_{jk}) > v_i\left(g^{FT}\right)$ and, conditional on g_{ij}, i becomes the hub on the path to global free trade in equilibrium.

Parts (i) and (ii) deal with "participation constraints". Part (i) says a country prefers to be an insider with a more attractive country and then a spoke on the path to global free trade over the permanent status quo of no agreements. Given Conditions 1 and 2, this must hold if $v_i\left(g_j^H\right) > v_i(\varnothing)$. Indeed, $v_i\left(g_j^H\right) > v_i(\varnothing)$ holds if the presence of tariff complementarity leads FTA formation to confer a positive externality on the outsider because then $v_i\left(g_j^H\right) > v_i(g_{jk}) > v_i(\varnothing)$. Part (ii) says the most attractive country i prefers to be an insider with the moderately attractive country j and then the hub on the path to global free trade over the permanent status quo of no agreements. Conditions 1 and 2 imply this must hold under symmetry and can can only fail under asymmetry if $v_i\left(g^{FT}\right) < v_i(\varnothing)$.

Part (iii) deals with the following situation for the least attractive countries j and k. On one hand, each such country could form an FTA with the most attractive country i and be an insider-turned-spoke on the path to global free trade. On the other hand, the two least attractive countries could form an FTA themselves but this would be a permanent FTA because the most attractive country i refuses to participate in any subsequent agreements

[14] Note that the absence of FTA exclusion incentives allows the non-member to benefit from tariff complementarity when FTA formation takes place at the insider–outsider network.

(i.e. $v_i(g_{jk}) > v_i(g^{FT}) > v_i(g_h^H)$ for $h \neq i$). Part (iii) says each of the least attractive countries prefers the former over the latter. In practice, $v_i(g_{jk}) > v_i(g^{FT})$ is likely when the most attractive country i is very attractive relative to the less attractive countries j and k which is exactly when being an insider-turned spoke with i is attractive compared to a permanent status quo of g_{jk}. Thus, part (iii) is fairly weak.

2.2 The general payoff properties and more trade models

Importantly, later results are driven by robust features of payoffs across numerous models. Lemma 1 formalizes this statement by relating the general payoff properties (Conditions 1-3 and Condition 4 in Section 5) to numerous trade models.[15] Supplemental Appendix A presents the alternative models and lemma proof (Supplemental Appendix B has all other proofs of results in the main text). Except for the political economy oligopolistic model, governments' payoff functions are national welfare and they set optimal tariffs.

Lemma 1. *Under symmetry, Conditions 1 and 3 are satisfied by (i) the political economy oligopolistic model and (ii) the competing importers model. Under asymmetry, there are ranges of the parameter spaces where Conditions 2-4 are satisfied by (i) the political economy oligopolistic model with market size asymmetry, (ii) the oligopolistic model with market size asymmetry, (iii) the competing exporters model with market size asymmetry and (iv) the competing importers model with either market size or technology asymmetry.*

3 Dynamic network formation games

3.1 Network transitions and preferences over transitions

The three country game, with $N = \{s, m, l\}$ denoting the set of countries, is very similar to Seidmann (2009). Like Seidmann (2009), I assume (i) at most one agreement (i.e. bilateral FTA or three country MFN agreement) can form in a period and (ii) agreements formed in previous periods cannot be severed. Thus, given the networks depicted in Figure 1, Table 1 illustrates the feasible network transitions within a period.[16] Hereafter, $g_{t-1} \to g_t$ denotes the feasible transition within the current period from g_{t-1} to g_t.

Having used backward induction to solve the equilibrium transitions in subsequent periods, players have preferences over current period feasible transitions. Given a network at

[15]Note, the FTA exclusion incentive fails to hold in the symmetric oligopoly and symmetric competing exporter models when the government objective function is national welfare and they set optimal tariffs.

[16]These transitions differ from Seidmann (2009) only because Seidmann's question of interest leads to an environment where countries can form CUs or FTAs.

Network at start of current period	Possible networks at end of current period
$\emptyset$	$\emptyset, g_{ij}, g_{ik}, g_{jk}, g^{FT}$
g_{ij}	$g_{ij}, g_i^H, g_j^H, g^{FT}$
g_i^H	g_i^H, g^{FT}
g^{FT}	g^{FT}

Table 1: Networks and feasible transitions within a period

the beginning of the current period g_{t-1} and a pair of transitions $g_{t-1} \to g_t$ and $g_{t-1} \to g'_t$, player i *prefers* g_t over g'_t if and only if $g_{t-1} \to g_t$ yields a strictly higher continuation payoff for player i than $g_{t-1} \to g'_t$. This preference is denoted $g_t \succ_i g'_t$. Further, g_t is *(strictly) most preferred* for country i in period t if $g_{t-1} \to g_t$ generates a (strictly) higher continuation payoff than any other transition $g_{t-1} \to g'_t$ where $g'_t \neq g_t$.

3.2 Actions, strategies and equilibrium concept

Each period can be characterized by the network g that exists at start of the period. Given an exogenous protocol specifying how countries make trade agreement proposals in a period, I refer to this "proposal game" as the subgame at network g (as in Seidmann (2009)).

I adopt a protocol where a *proposer* country proposes a trade agreement and the proposed members, i.e. *recipients*, then respond by accepting or not accepting. In each period, country l is the first proposer (stage 1), followed by country m (stage 2) and then country s (stage 3). If each recipient country accepts the proposal in a given stage, the proposed agreement forms and the period ends. But, if at least one of the recipient countries rejects the proposal, or the proposer makes no proposal, then the protocol moves to the subsequent stage. Thus, the period ends after either (i) an agreement forms or (ii) no agreement forms despite each country having the opportunity to be the proposer.

As the proposer, a country can propose an agreement that has not yet formed and to which it will be a member. In the "bilateralism game", Table 2 illustrates the available proposals for each country i and for each subgame at network g with $P_i(g)$ denoting the set of such proposals and $\rho_i(g) \in P_i(g)$ denoting a proposal. In Table 2, ij denotes the FTA between i and j, FT denotes the three country MFN agreement that takes the world to global free trade, and ϕ denotes the proposer elects to make no proposal. In the "multilateralism game", the only possible agreement is the three country MFN agreement taking the world to global free trade. Thus, the game essentially reduces to a single period game with $P_i(\emptyset) = \{\phi, FT\}$ for each i. Upon receiving a proposal $\rho_i(g)$, each recipient country j (i.e. a country of the proposed agreement) responds by announcing $r_j(g, \rho_i(g)) \in \{Y, N\}$ where Y (N) denotes j accepts (does not accept) the proposal.

	$P_i(g)$	$P_j(g)$	$P_k(g)$
$\emptyset$	$\{\phi, ij, ik, FT\}$	$\{\phi, ij, jk, FT\}$	$\{\phi, ik, jk, FT\}$
g_{ij}	$\{\phi, ik, FT\}$	$\{\phi, jk, FT\}$	$\{\phi, ik, jk, FT\}$
g_i^H	$\{\phi, FT\}$	$\{\phi, jk, FT\}$	$\{\phi, jk, FT\}$
g^{FT}	$\{\phi\}$	$\{\phi\}$	$\{\phi\}$

Table 2: Proposer country's action space for each subgame in the bilateralism game

Given the protocol, country i's Markov strategy in the bilateralism game must do two things for every subgame at network g: (i) assign a proposal $\rho_i(g) \in P_i(g)$ for the stage where country i is the proposer and (ii) assign a response $r_i(g, \rho_j(g)) \in \{Y, N\}$ to any proposal country i may receive from some other country $j \neq i$. I now use backward induction to solve for a pure strategy Markov Perfect Equilibrium.[17]

4 Symmetric countries

4.1 Bilateralism game

To begin the backward induction with symmetric countries (i.e. $\alpha_l = \alpha_m = \alpha_s$), consider a subgame at a hub–spoke network $g = g_i^H$. Since each FTA is mutually beneficial by Condition 1, spokes form their own FTA.

Lemma 2. *Consider the subgame at a hub-spoke network g_i^H and suppose $v_h\left(g^{FT}\right) > v_h\left(g_i^H\right)$ for $h \neq i$. Then, spokes form their own FTA and global free trade is attained (i.e. $g_i^H \rightarrow g^{FT}$).*

Now roll back to a subgame at an insider–outsider network $g = g_{ij}$. Here, insiders face a trade off. Myopically, becoming the hub is attractive due to reciprocal preferential access exchanged with the outsider: $v_i\left(g_i^H\right) > v_i(g_{ij})$. However, the would–be hub anticipates the subsequent spoke-spoke FTA erodes the value of reciprocal preferential access enjoyed as the hub with *each* spoke country. Indeed, the degree of preference erosion is sufficiently large that insiders have an FTA exclusion incentive: $v_i(g_{ij}) > v_i\left(g^{FT}\right)$. Thus, an insider i wants to become the hub rather than remain a permanent insider with j if and only if

[17] For convenience, I make two assumptions that restrict attention to certain Markov Perfect Equilibria. First, given the simultaneity of responses to a proposal for a three country MFN agreement, I assume countries respond to such proposals affirmatively if they prefer global free trade over the status quo. That is, $r_i(g, FT) = Y$ if $g^{FT} \succ_i g$ in the subgame at network g. I also assume a recipient country responds with $r_i(g, \rho_j(g)) = Y$ when responding with $r_i(g, \rho_j(g)) = N$ would merely delay formation of the proposed agreement to a later stage of the current period. This can be motivated by the presence of an arbitrarily small cost involved in making a response.

$v_i\left(g_i^H\right)+\frac{\delta}{1-\delta}v_i\left(g^{FT}\right)>\frac{1}{1-\delta}v_i\left(g_{ij}\right)$ which reduces to the No Exclusion (NE) condition:

$$\delta<\bar{\delta}_{i,j}^{NE}\left(\alpha\right)\equiv\frac{v_i\left(g_i^H\right)-v_i\left(g_{ij}\right)}{v_i\left(g_i^H\right)-v_i\left(g^{FT}\right)}=\frac{v_i\left(g_i^H\right)-v_i\left(g_{ij}\right)}{\left[v_i\left(g_i^H\right)-v_i\left(g_{ij}\right)\right]+\left[v_i\left(g_{ij}\right)-v_i\left(g^{FT}\right)\right]} \tag{2}$$

where $\alpha\equiv(\alpha_s,\alpha_m,\alpha_l)$ and, given symmetry, $\bar{\delta}^{NE}\equiv\bar{\delta}_{i,j}^{NE}\left(\alpha\right)$. When an insider's No Exclusion condition holds (fails) then $\delta<(>)\bar{\delta}^{NE}$ and the myopic attractiveness of becoming the hub dominates (is dominated by) the subsequent preference erosion. Thus, an insider wants to become the hub (remain an insider forever). Lemma 3 formalizes the role of the No Exclusion condition.

Lemma 3. *Suppose Condition 1 holds and consider a subgame at an insider–outsider network g_{ij}. The equilibrium outcomes of the subgame are: (i) no agreement (i.e. $g_{ij}\to g_{ij}$) when $\delta>\bar{\delta}^{NE}$, (ii) an FTA between the outsider and either insider (i.e. $g_{ij}\to g_i^H$ and $g_{ij}\to g_j^H$) when $\delta<\bar{\delta}^{NE}$ and the outsider is the first proposer and (iii) an FTA between the outsider and the first insider in the protocol (i.e. $g_{ij}\to g_i^H$) when $\delta<\bar{\delta}^{NE}$ and the outsider is not the first proposer.*

When the No Exclusion condition is violated, $\delta>\bar{\delta}^{NE}$, each insider prefers remaining an insider over becoming the hub on the path to global free trade. Regardless of an insider's position in the protocol, it anticipates the other insider will reject any future proposal from the outsider. In turn, each insider refrains from making a proposal. Thus, the mutual fear of preference erosion allows insiders to remain insiders when $\delta>\bar{\delta}^{NE}$.

However, each insider wants to become the hub when the No Exclusion condition holds, $\delta<\bar{\delta}^{NE}$, because the fear of preference erosion is sufficiently small. Thus, while the FTA exclusion incentive implies each insider would reject a proposed move directly to global free trade, each insider wants to form an FTA with the outsider and thereby enjoy the hub benefits of preferential access with each spoke country on the path to global free trade. However, which hub-spoke network(s) emerge in equilibrium depends on the outsider's position in the protocol. If an insider i is the first proposer, it proposes an FTA with the outsider k who accepts and the hub-spoke network g_i^H emerges. But, if the outsider is the first proposer then its indifference regarding the identity of its partner, and the fact that either insider will accept an FTA proposal, generates multiplicity. Now an FTA between the outsider and either insider is an equilibrium outcome and thus either hub-spoke network, g_i^H or g_j^H, could emerge. Nevertheless, $\delta<\bar{\delta}^{NE}$ implies the fear of preference erosion is sufficiently small and some hub-spoke network emerges in the subgame at an insider-outsider network.

Now roll back to the subgame at the empty network $g=\emptyset$. Solving the equilibrium outcome in this subgame reveals the equilibrium path of networks. To do so, define $\bar{\delta}$ such that $v_i(g_{ij}) + \delta v_i(g_j^H) + \frac{\delta^2}{1-\delta} v_i(g^{FT}) < \frac{1}{1-\delta} v_i(g^{FT})$ if and only if $\delta > \bar{\delta}$.[18] That is, $\delta > \bar{\delta}$ implies a country prefers a direct move to global free trade over being an insider–turned–spoke on the path to global free trade. Proposition 1 now follows, remembering the protocol specifies country l as the first proposer followed by country m and then country s.

Proposition 1. *Suppose Conditions 1 and 3 hold. The equilibrium path of networks is (i) g_{sl} or g_{ml} when $\delta > \bar{\delta}^{NE}$, (ii) g^{FT} when $\delta \in (\bar{\delta}, \bar{\delta}^{NE})$, and (iii) $g_{sl} \to g_l^H \to g^{FT}$ when $\delta < \bar{\delta}^{NE}$.*

When the No Exclusion condition is violated, $\delta > \bar{\delta}^{NE}$, the mutual fear of preference erosion is sufficiently large that remaining insiders is strictly most preferred for any pair of insiders. However, either insider–outsider network g_{sl} or g_{ml} can emerge because symmetry creates indifference on the part of country l, the first proposer, regarding its FTA partner.

When the No Exclusion condition holds, any bilateral FTA eventually leads to global free trade via a hub-spoke network. For the insider-turned-spoke, they face a trade-off between this path and a direct move to global free trade. While the FTA exclusion incentive makes being an insider-turned-spoke attractive, a direct move to global free trade eliminates the discrimination faced as a spoke. When $\delta \in (\bar{\delta}, \bar{\delta}^{NE})$, the discrimination faced as a spoke dominates the FTA exclusion incentive and a country prefers a direct move to global free trade. Because country s is the third proposer in the protocol and thus can never be the hub in equilibrium (see Lemma 3), it proposes global free trade when $\delta \in (\bar{\delta}, \bar{\delta}^{NE})$ knowing that the other countries will accept given $v_h(g^{FT}) > v_h(\emptyset)$ for any h. In turn, any country receiving a proposal that results in becoming an insider-turned-spoke will reject the proposal. Thus, a direct move to global free trade emerges when $\delta \in (\bar{\delta}, \bar{\delta}^{NE})$.

Once $\delta < \bar{\delta}$, the FTA exclusion incentive dominates any discrimination faced as a spoke. Thus, a country prefers being an insider-turned-spoke on the path to global free trade over a direct move to global free trade. Hence, country l proposes an FTA with country s and country s accepts. Country s accepts knowing it will never be the hub (Lemma 3). Further, country l does not propose an FTA with country m knowing m will reject the proposal so it can then propose an FTA with s and be the insider-turned-hub on the path to global free trade. Thus, $g_{sl} \to g_l^H \to g^{FT}$ is the unique equilibrium path of networks once $\delta < \bar{\delta}$.

4.2 Role of FTAs under symmetry

To isolate the role of FTAs, I follow Saggi and Yildiz (2010, 2011) by comparing the equilibrium outcome of (i) the "bilateralism game" of the previous section and (ii) the "multilater-

[18] Simple manipulation reveals $\bar{\delta} = \frac{v_i(g_{ij}) - v_i(g^{FT})}{v_i(g^{FT}) - v_i(g_j^H)}$.

alism game" which removes the possibility of FTAs. FTAs are strong building (stumbling) blocs when global free trade is only attained in the bilateralism (multilateralism) game.

Since the only possible agreement in the multilateralism game is the three country MFN agreement, each country has veto power and the equilibrium characterization is simple.

Proposition 2. *Let $\alpha_l \geq \alpha_m \geq \alpha_s$ in the multilateralism game. The equilibrium path of networks is (i) a direct move to global free trade (i.e. $\emptyset \rightarrow g^{FT}$) if $v_i\left(g^{FT}\right) > v_i(\emptyset)$ for all i, but (ii) the empty network (i.e. $\emptyset \rightarrow \emptyset$) if $v_i\left(g^{FT}\right) < v_i(\emptyset)$ for some i.*

Given Condition 1, an immediate implication of Proposition 2 is that global free trade is the unique equilibrium path of networks under symmetry.

Corollary 1 now follows from Propositions 1 and 2 and summarizes the role of FTAs.

Corollary 1. *Suppose countries are symmetric. Then, FTAs are strong stumbling blocs when $\delta > \bar{\delta}^{NE}$. However, global free trade is attained in the bilateralism and multilateralism games when $\delta < \bar{\delta}^{NE}$.*

Corollary 1 emphasizes the destructive role of FTAs caused by insiders' fear of preference erosion. Under symmetry, no country vetoes global tariff elimination when non-discriminatory liberalization is the only form of liberalization. However, the opportunity to form discriminatory FTAs leads to a single FTA when fear of preference of erosion is sufficiently strong (i.e. $\delta > \bar{\delta}^{NE}$). Thus, FTAs are strong stumbling blocs when $\delta > \bar{\delta}^{NE}$.

Corollary 1 is a strong result given FTAs can be strong stumbling blocs under symmetry. Saggi and Yildiz (2010, 2011) find FTAs are never strong stumbling blocs and, under symmetry, the presence of FTAs leads to global free trade. Moreover, Saggi et al. (2013) find Customs Unions (CUs) can be strong stumbling blocs only when countries are sufficiently asymmetric. Thus, my strong stumbling bloc result under symmetry emphasizes the dynamic role played by preference erosion gives a fundamentally different mechanism for the destructive role of preferential trade agreements (i.e. FTAs or CUs) than Saggi et al. (2013).

5 Asymmetric countries

5.1 Bilateralism game

I now use backward induction to solve the equilibrium path of networks with asymmetric countries where s, m and l denote the "small", "medium" and "large" countries and, hence, $\alpha_s < \alpha_m < \alpha_l$.

To begin, note that, like the symmetric case, hub-spoke networks expand to global free trade because FTAs mutually benefit spokes. However, asymmetry creates three important differences in subgames at insider–outsider networks.

First, the strength of an insider's FTA exclusion incentive depends on the characteristics of itself and its insider partner. Hence, each insider has a distinct No Exclusion condition and $\bar{\delta}_{i,j}^{NE}(\alpha)$ no longer reduces to $\bar{\delta}^{NE}$. In turn, the eventual emergence of global free trade from subgames at an insider-outsider networks can depend on the insiders' identity.

Second, a larger insider may engage in FTA formation with the outsider, and thereby become the hub, merely to avoid becoming a spoke. To see this, suppose (i) s is the outsider and willing to form an FTA with either insider m or l, but (ii) $\delta \in \left(\bar{\delta}_{l,m}^{NE}(\alpha), \bar{\delta}_{m,l}^{NE}(\alpha)\right)$ so that m wants to become the hub even though l ideally wants to remain a permanent insider. Then, given global free trade emerges from any hub-spoke network and s prefers FTA formation with a larger country, the anticipation of being discriminated against as a spoke induces l to become the hub by proposing an FTA with s.

Third, an outsider may refuse FTA formation with an insider. Given spokes form their own FTA, an outsider i prefers forming an FTA with an insider j and becoming a spoke rather than remaining a permanent outsider if and only if $v_i\left(g_j^H\right) + \frac{\delta}{1-\delta} v_i\left(g^{FT}\right) > \frac{1}{1-\delta} v_i(g_{jk})$. This reduces to the Free Trade–Outsider (FT–O) condition:

$$\delta > \bar{\delta}_{i,j}^{FT-O}(\alpha) \equiv \frac{v_i(g_{jk}) - v_i\left(g_j^H\right)}{v_i\left(g^{FT}\right) - v_i\left(g_j^H\right)}. \tag{3}$$

Given $v_i\left(g^{FT}\right) > v_i\left(g_j^H\right)$, then $\bar{\delta}_{i,j}^{FT-O}(\alpha) \in (0,1)$ if and only if $v_i\left(g^{FT}\right) > v_i(g_{jk}) > v_i\left(g_j^H\right)$. In this case, an outsider faces a tension between the *myopic* incentive to resist becoming a spoke and the *future* appeal of global free trade. Thus, an outsider i wants to become the spoke by forming an FTA with j when $\delta > \bar{\delta}_{i,j}^{FT-O}(\alpha)$ but refuses FTA formation with the insider j when $\delta < \bar{\delta}_{i,j}^{FT-O}(\alpha)$. Indeed, once $\delta < \min\left\{\bar{\delta}_{i,j}^{FT-O}(\alpha), \bar{\delta}_{i,k}^{FT-O}(\alpha)\right\} < 1$ then only a direct move to global free trade can induce the outsider's participation in liberalization.

Under symmetry, the FT–O condition was irrelevant (i.e. $\bar{\delta}_{i,j}^{FT-O}(\alpha) < 0$) because all FTAs were mutually beneficial. However, part (ii) of Condition 2 says an outsider may have a myopic incentive to refuse an FTA with a country smaller than itself. Thus, l may refuse an FTA with m and/or s in the subgame at g_{sm} while m may refuse an FTA with s in the subgame at g_{sl}. Moreover, Condition 2 implies l may refuse participation in *any* subsequent agreement whether it be an FTA or a direct move to global free trade (this happens when $v_l(g_{sm}) > v_l\left(g^{FT}\right) > v_l\left(g_m^H\right)$ and hence $\bar{\delta}_{l,m}^{FT-O}(\alpha) > 1$).

Rolling back to the subgame at the empty network, Proposition characterizes the equilibrium path of networks when the No Exclusion condition is violated for the two largest countries. To this end, let $\hat{\delta}_{i,j}^{NE}(\alpha) \equiv \max\left\{\bar{\delta}_{i,j}^{NE}(\alpha), \bar{\delta}_{j,i}^{NE}(\alpha)\right\}$ and remember $g \succ_i g'$ is a comparison of *continuation payoffs* resulting from transitions to g and g'.

Proposition 3. *Suppose Conditions 2-3 hold and let $\delta > \hat{\delta}_{m,l}^{NE}(\alpha)$. Then, global free trade does not emerge in equilibrium. The equilibrium path of networks is g_{ml} unless $v_l(g_{sm}) > v_l(g_{ml})$ and $\emptyset \succ_l g$ for $g = g_{sl}, g^{FT}$ in which case the equilibrium path of networks is g_{sm}.*

Proposition 3 emphasizes that, like Proposition 1 under symmetry, No Exclusion conditions remain crucial for determining whether global free trade emerges in the presence of FTAs.

When $\delta > \hat{\delta}_{m,l}^{NE}(\alpha)$ and m and l are insiders, the mutual fear of preference erosion allows m and l to refrain from making proposals and, hence, remain permanent insiders. Indeed, $\delta > \hat{\delta}_{m,l}^{NE}(\alpha)$ implies that, in the subgame at the empty network, this is strictly most preferred for m regardless of the outcomes in subgames at the insider-outsider networks g_{sl} and g_{sm}. Moreover, Condition 2 implies the same is true for l when $v_l(g_{ml}) > v_l(g_{sm})$. Thus, in these cases, the equilibrium path of networks is the permanent FTA between m and l, g_{ml}.

However, given l's FTA exclusion incentive, $v_l(g_{sm}) > v_l(g_{ml})$ implies $v_l(g_{sm}) > v_l\left(g^{FT}\right)$ meaning l refuses *any* subsequent agreement as the outsider. Thus, s and m remain permanent insiders conditional on becoming insiders. Moreover, $v_l(g_{sm}) > v_l(g_{ml})$ implies free riding on this permanent FTA is strictly most preferred for l in the subgame at the empty network. Nevertheless, s is the third proposer and Conditions 2-3 imply s prefers a direct move to global free trade or an FTA with l over a permanent FTA with m. Thus, the equilibrium path of networks is g_{sm} only if, as the outsider, l credibly refuses proposals from s for an FTA and the three country MFN agreement; otherwise, g_{ml} again emerges.

Determining if global free trade eventuates in equilibrium when $\delta < \hat{\delta}_{m,l}^{NE}(\alpha)$ requires knowing if global free trade eventuates conditional on reaching each insider-outsider network. This generally depends on the interplay between the six No Exclusion conditions and the three relevant Free Trade-Outsider conditions ($\bar{\delta}_{l,m}^{FT-O}(\alpha)$, $\bar{\delta}_{l,s}^{FT-O}(\alpha)$ and $\bar{\delta}_{m,s}^{FT-O}(\alpha)$), producing numerous possible combinations of outcomes. Hence, Condition 4 restricts the relationship between the various No Exclusion and Free Trade-Outsider conditions (Lemma 1 establishes Conditions 2-4 jointly hold in numerous trade models).

Condition 4. (i) $\hat{\delta}_{m,l}^{NE}(\alpha) < \min\left\{\hat{\delta}_{s,l}^{NE}(\alpha), \bar{\delta}_{m,s}^{NE}(\alpha)\right\}$

(ii) $\min\left\{\hat{\delta}_{m,l}^{NE}(\alpha), \bar{\delta}_{m,s}^{FT-O}(\alpha)\right\} < \bar{\delta}_{l,s}^{NE}(\alpha)$

(iii) $\frac{\partial \hat{\delta}_{m,l}^{NE}(\alpha)}{\partial \alpha_{hs}} < 0$ for $h = m, l$ or $\frac{\partial \hat{\delta}_{m,l}^{NE}(\alpha)}{\partial \alpha_{ms}} + \frac{\partial \hat{\delta}_{m,l}^{NE}(\alpha)}{\partial \alpha_{ls}} < 0$.

Part (i) essentially says the largest insiders, m and l, have the tightest No Exclusion conditions. This is intuitive given (2): the relatively low attractiveness of the outsider s strengthens the FTA exclusion incentives of m and l (i.e. raises $v_i(g_{ij}) - v_i\left(g^{FT}\right)$) and weakens the appeal of becoming the hub via an FTA with the outsider s (i.e. lowers $v_i\left(g_i^H\right)-v_i\left(g_{ij}\right)$). Importantly, once $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$, part (i) says global free trade emerges from an insider-outsider network if the outsider wants to form an FTA with either insider.

However, as discussed above, an outsider may refuse FTAs with both insiders. In the subgame at g_{sl}, Condition 2 says the outsider m may not want to form an FTA with s. This scuttles FTA formation if, in addition, l does not want to form an FTA with m (i.e. $\delta<\bar{\delta}_{l,s}^{NE}(\alpha)$). Part (ii) of Condition 4 rules this possibility out once $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$.[19] Thus, the equilibrium path of networks conditional on s and l being insiders is either $g_{sl}\to g_l^H\to g^{FT}$ or $g_{sl}\to g^{FT}$ once $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$. In turn, given the equilibrium path of networks conditional on m and l being insiders is $g_{ml}\to g_l^H\to g^{FT}$, the only insider-outsider network that may fail to eventually reach global free trade is g_{sm}. Moreover, such failure can only occur if, as the outsider, l refuses *any* subsequent agreement.

Part (iii) has the spirit of part (i) but is more specific and only deals with the No Exclusion conditions of m and l as insiders: greater asymmetry tightens the slackest No Exclusion condition where greater asymmetry means either (i) rising α_{ls} *or* α_{ms} or (ii) a simultaneous marginal increase in α_{ls} *and* α_{ms}. Part (iii) allows Section 5.2 to investigate how greater asymmetry affects the strong building bloc-strong stumbling bloc analysis.

With Condition 4 in place, the following proposition characterizes the equilibrium emergence of global free trade when $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$.

Proposition 4. *Suppose Conditions 2-4 hold and $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$. Then global free trade emerges on any equilibrium path of networks unless $g_{sm}\succ_l g_{ml}$ and $\emptyset\succ_l g$ for $g=g_{sl},g^{FT}$ in which case the equilibrium path of networks is g_{sm}.*

While Proposition 4 has a similar flavor to Proposition 1 under symmetry, $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$ is necessary but not sufficient for the emergence of global free trade.

How can global free trade fail to emerge in equilibrium once $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$? Some agreement must form in equilibrium given m and l mutually benefit, relative to no agreements, from their own FTA (see Condition 3). But, the only insider-outsider network that may not eventually reach global free trade is g_{sm} which can only happen if, as the outsider, l refuses participation in any subsequent agreement. Thus, the permanent FTA between s and m emerges in equilibrium if: (i) as the outsider, l credibly refuses *any* proposal (i.e. FTA and three country MFN agreement) and (ii) l prefers free riding on this permanent FTA over the reciprocal preferential market access enjoyed as an insider with m and then the hub.

Before revisiting the strong building–strong stumbling bloc analysis, I analyze the path

[19] That is, $\delta<\hat{\delta}_{m,l}^{NE}(\alpha)$ implies $\delta<\bar{\delta}_{l,s}^{NE}(\alpha)$ whenever $\delta<\bar{\delta}_{m,s}^{FT-O}(\alpha)$. Moreover, part (ii) of Condition 4 must hold if $\hat{\delta}_{m,l}^{NE}(\alpha)=\bar{\delta}_{l,m}^{NE}(\alpha)$ because Condition 2 implies $\bar{\delta}_{l,m}^{NE}(\alpha)<\bar{\delta}_{l,s}^{NE}(\alpha)$.

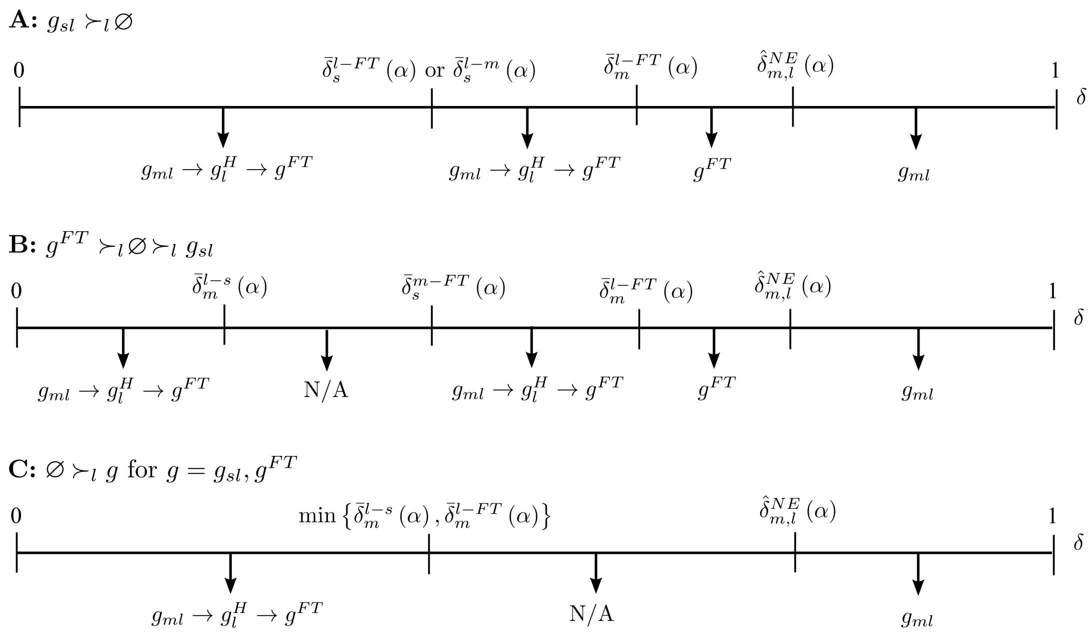

Figure 2: Equilibrium path of networks when $g_{ml} \succ_l g_{sm}$

of FTAs leading to global free trade. In particular, Proposition 5 characterizes conditions where $g_{ml} \to g_l^H \to g^{FT}$ is the unique equilibrium path of networks and is depicted in Figure 2. These conditions are sufficient rather than necessary and, thus, merely intended to convey the pervasiveness of $g_{ml} \to g_l^H \to g^{FT}$.[20]

Proposition 5. *Suppose Conditions 2-4 hold and FTAs emerge on an equilibrium path of networks leading to global free trade. Further suppose that either (i) $g_{sl} \succ_l \varnothing$, or (ii) $g^{FT} \succ_l \varnothing \succ_l g_{sl}$ and either $g^{FT} \succ_s g_{sm}$ or $g_{ml} \succ_m g_{sm}$, or (iii) $\varnothing \succ_l g$ for $g = g_{sl}, g^{FT}$ and $g_{ml} \succ_m g$ for $g = g_{sm}, g^{FT}$. Then, the unique equilibrium path of networks is $g_{ml} \to g_l^H \to g^{FT}$.*

Importantly, regardless of the particular panel, s prefers either an FTA with l or a direct move to global free trade over an FTA with m once $\delta < \hat{\delta}_{m,l}^{NE}(\alpha)$.[21] Thus, ideally, s wants to avoid proposing $\rho_s(\phi) = sm$ in stage 3. When condition (i) of Proposition 5 holds, Panel A of Figure 2 is relevant. What will s propose? Relative to proposing $\rho_s(\varnothing) = sm$ or $\rho_s(\varnothing) = FT$, the *myopic* attractiveness of an FTA with l makes $\rho_s(\varnothing) = sl$ appealing. Conversely, to avoid *future* discrimination as a spoke after an FTA with l, it is appealing to propose $\rho_s(\varnothing) = FT$ or, if g_{sm} expands directly to global free trade, $\rho_s(\varnothing) = sm$. Thus, s proposes (does not propose) $\rho_s(\varnothing) = sl$ when δ falls below (exceeds) a threshold. In particular, if l will accept $\rho_s(\varnothing) = FT$ (i.e. $g^{FT} \succ_l \varnothing$), s proposes $\rho_s(\varnothing) = FT$ when

[20] The segments labeled N/A in Figure 2 indicate Proposition 5 does not specify the equilibrium.

[21] Intuitively, (i) $g_{sm} \succ_s g_{sl}$ only if $g_{sm} \to g^{FT}$ which then implies $g^{FT} \succ_s g_{sm}$ and (ii) $g_{sm} \succ_s g^{FT}$ only if $g_{sm} \to g_m^H \to g^{FT}$ but then $g_{sl} \succ_s g_{sm}$.

$\delta > \bar{\delta}_s^{l-FT}(\alpha)$ but $\rho_s(\emptyset) = sl$ when $\delta < \bar{\delta}_s^{l-FT}(\alpha)$. Conversely, if l refuses $\rho_s(\emptyset) = FT$ (i.e. $\emptyset \succ_l g^{FT}$), s proposes $\rho_s(\emptyset) = sm$ when $\delta > \bar{\delta}_s^{l-m}(\alpha)$, which then expands directly to global free trade, but $\rho_s(\emptyset) = sl$ when $\delta < \bar{\delta}_s^{l-m}(\alpha)$.

When s proposes $\rho_s(\emptyset) = sm$ or $\rho_s(\emptyset) = FT$ then m faces a trade-off: propose $\rho_m(\emptyset) = FT$ or $\rho_m(\emptyset) = ml$.[22] The former eliminates the *future* discrimination faced as a spoke but the latter provides the *myopic* benefit of reciprocal preferential access with l. Thus, m proposes (in stage 2) and accepts (in stage 1) global free trade when $\delta > \bar{\delta}_m^{l-FT}(\alpha)$ but the FTA with l when $\delta < \bar{\delta}_m^{l-FT}(\alpha)$. Further, the prospect of future discrimination as a spoke when s proposes $\rho_s(\emptyset) = sl$ (i.e. $\delta < \bar{\delta}_s^{l-FT}(\alpha)$ or $\delta < \bar{\delta}_s^{l-m}(\alpha)$) induces m to propose/accept an FTA with l.

The analysis is similar in panel B of Figure 2 which is relevant when condition (ii) holds. Given l refuses s's proposal of $\rho_s(\emptyset) = sl$ in stage 3, s proposes either $\rho_s(\emptyset) = sm$ or $\rho_s(\emptyset) = FT$. Similar to the previous paragraph, s proposes $\rho_s(\emptyset) = FT$ when δ exceeds the critical value $\bar{\delta}_s^{m-FT}(\alpha)$ to avoid future discrimination as a spoke, but proposes $\rho_s(\emptyset) = sm$ when $\delta < \bar{\delta}_s^{m-FT}(\alpha)$. In turn, $\delta > \bar{\delta}_s^{m-FT}(\alpha)$ implies m faces the trade off in the previous paragraph between proposing/accepting an FTA with l or a direct move to global free trade with the answer again depending on $\delta \gtrless \bar{\delta}_m^{l-FT}(\alpha)$. Moreover, the FTA between m and l may still emerge in equilibrium when $\rho_s(\emptyset) = sm$. As above, $\rho_s(\emptyset) = sm$ only if $g_{sm} \to g^{FT}$ which implies l will accept an FTA proposal from m. Thus, $g_{ml} \to g_l^H \to g^{FT}$ again emerges if $g_{ml} \succ_m g_{sm}$ which is true when δ falls below the threshold $\bar{\delta}_m^{l-s}(\alpha)$ because then m places sufficient weight on the myopic attractiveness of an FTA with l rather than s.

The analysis is again similar in Panel C of Figure 2 which is relevant for condition (iii). Here, l refuses $\rho_s(\emptyset) = sl$ and $\rho_s(\emptyset) = FT$, forcing s to propose $\rho_s(\emptyset) = sm$. However, if l prefers an FTA with m over becoming the outsider (i.e. $g_{ml} \succ_l g_{sm}$), m proposes/accepts an FTA with l when $\delta < \min\left\{\bar{\delta}_m^{l-s}(\alpha), \bar{\delta}_m^{l-FT}(\alpha)\right\}$ to exploit the myopic attractiveness of an FTA with l rather than an FTA with s or a direct move to global free trade.

The following section now revisits the role of FTAs when countries are asymmetric.

5.2 Role of FTAs under asymmetry

To begin, Lemma 4 explicitly states how greater asymmetry affects $\hat{\delta}_{m,l}^{NE}(\alpha)$. Put simply, greater asymmetry strengthens the fear of preference erosion which increases (decreases) the extent that m and l remain insiders (global free trade is attained).

[22] Note, $v_i\left(g^{FT}\right) > v_i(g_{sm})$ for all i must hold in the case where s proposes $\rho_m(\emptyset) = sm$ and $g_{sm} \to g^{FT}$. As such, in this case, $g^{FT} \succ_i g_{sm}$ for all i.

Lemma 4. *Suppose Conditions 2-4 hold. Then greater asymmetry increases the extent to which m and l remain permanent insiders in equilibrium and reduces the extent to which global free trade is eventually attained in equilibrium. If $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{hs}} < 0$ for $h = m, l$ then greater asymmetry via a higher α_{ls} or α_{ms} reduces $\hat{\delta}^{NE}_{ml}(\alpha)$. If $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{sm}} + \frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{ls}} < 0$ then greater asymmetry via a simultaneous marginal increase in α_{ls} and α_{ms} reduces $\hat{\delta}^{NE}_{ml}(\alpha)$.*

Corollary 2, following directly from Propositions 3-4, now summarizes how the role of FTAs depends on asymmetry and is the central result of the paper. Note that part (iv) of Condition 2 implies $v_l\left(g^{FT}\right) > v_l(\emptyset)$ reduces to $\alpha_{ms} > \bar{\alpha}_{ms}(\alpha)$ where $\frac{\partial \bar{\alpha}_{ms}(\cdot)}{\partial \alpha_{ls}} > 0$.

Corollary 2. *Suppose Conditions 2-4 hold. FTAs are strong stumbling blocs when l and m are sufficiently symmetric but m and s are sufficiently asymmetric (i.e. two "larger" and one "smaller" country): $\alpha_{ms} > \bar{\alpha}_{ms}(\alpha)$ and $\delta > \hat{\delta}^{NE}_{m,l}(\alpha)$. FTAs are strong building blocs when l and m are sufficiently asymmetric but m and s are sufficiently symmetric (i.e. one "larger" and two "smaller" countries): $\alpha_{ms} < \bar{\alpha}_{ms}(\alpha)$ and either (i) $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$ if $g_{ml} \succ_l g_{sm}$ or (ii) $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$ and $g_{sl} \succ_l \emptyset$ if $g_{sm} \succ_l g_{ml}$.*

Lemma 1 implies Corollary 2 is robust to various trade models. But, for illustration, Figure 3 depicts Corollary 2 using the political economy oligopolistic model with an exogenous common tariff $\tau = \frac{1}{4}\alpha_s$ and δ fixed at $\bar{\delta} \equiv .47$. In Figure 3, $v_l(g_{ml}) > v_l(g_{sm})$ and thus global free trade emerges in the bilateralism game if and only if $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$.[23]

To begin, consider the multilateralism game. By Proposition 2, g^{FT} is the unique equilibrium path of networks in the band between the $\alpha_{ms} = \alpha_{ls}$ and $\alpha_{ms} = \bar{\alpha}_{ms}(\alpha)$ lines. That is, l does not block global free trade when l and m are sufficiently similar in size because the world market is attractive enough that the market access received compensates for the domestic market access given up. Outside the band, l blocks global free trade so the empty network is the unique equilibrium path of networks. That is, when m and l are sufficiently different in size, the world market is so small that the market access received by l does not compensate for the domestic market access given up.

Now consider the bilateralism game. Given $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{hs}} < 0$ for $h = m, l$, the downward sloping bold line is a contour curve with $\hat{\delta}^{NE}_{m,l}(\alpha)$ constant.[24] In Figure 3, $\hat{\delta}^{NE}_{m,l}(\alpha) = \bar{\delta} \equiv .47$. Moreover, higher contour curves represent a lower $\hat{\delta}^{NE}_{m,l}(\alpha)$ because greater asymmetry

[23] $\alpha_{ls} < 1.75$ ensures Condition 2 holds. $\alpha_{ms} > \underline{\alpha}_{ms}(\alpha)$ ensures $v_l\left(g^H_m\right) > v_l(g_{sm})$ which renders part (iii) of Condition 3 irrelevant and, together with $v_l(g_{ml}) > v_l\left(g^{FT}\right)$, also implies $v_l(g_{ml}) > v_l(g_{sm})$.

[24] As shown in the proof of Lemma 1, $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{hs}} < 0$ for $h = m, l$ is true of all trade models therein except the competing importers model where $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{ls}} < 0$ but $\frac{\partial \hat{\delta}^{NE}_{m,l}(\alpha)}{\partial \alpha_{ms}} > 0$. But, the same economic intuition still applies except rising asymmetry must be either a rise in α_{ls} only or a joint increase in α_{ls} and α_{ms}. Graphically, the $\hat{\delta}^{NE}_{m,l}(\alpha)$ contour curve is upward sloping and so the interpretation of above (below) the Figure 3 contour curve would apply to the right (left) of the contour curve.

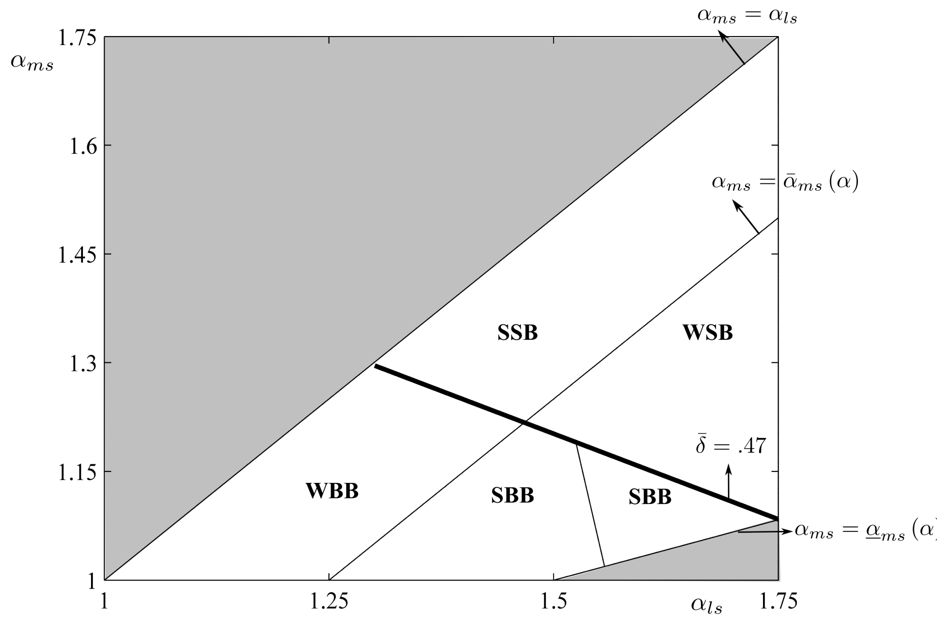

Figure 3: Role of FTAs under asymmetry when $\delta = .47$

reduces $\hat{\delta}^{NE}_{m,l}(\alpha)$. Hence, $\bar{\delta} = .47 \underset{(<)}{>} \hat{\delta}^{NE}_{m,l}(\alpha)$ above (below) the contour curve in Figure 3. Propositions 3-4 imply, given $v_l(g_{ml}) > v_l(g_{sm})$, global free trade is attained if and only if $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$. Thus, given $\delta = .47$, global free trade is attained below, but not above, the $\bar{\delta} = .47$ contour curve. That is, global free trade is not attained (is attained) in the bilateralism game when m and l are sufficiently larger than s (sufficiently similar to s). Intuitively, by strengthening preference erosion fears, greater asymmetry via a higher α_{ms} and α_{ls} increases the value of reciprocal preferential market access protected by m and l as insiders. The strong building–strong stumbling bloc dichotomy now emerges easily.

FTAs are strong stumbling blocs (SSB) when global free trade is attained in the multilateralism but not the bilateralism game. This is the area above the $\bar{\delta} = .47$ contour curve and inside the band between the $\alpha_{ms} = \alpha_{ls}$ and $\alpha_{ms} = \bar{\alpha}_{ms}(\alpha)$ lines. Here, m is sufficiently larger than s (i.e. above the contour curve) while m and l are sufficiently similar (i.e. inside the band). Thus, FTAs are strong stumbling blocs with two "larger" and one "smaller" country. Conversely, FTAs are strong building blocs (SBB) when global free trade is attained in the bilateralism but not the multilateralism game which is the area below the $\bar{\delta} = .47$ contour curve and outside the band between the $\alpha_{ms} = \alpha_{ls}$ and $\alpha_{ms} = \bar{\alpha}_{ms}(\alpha)$ lines. Here, m is sufficiently similar to s (i.e. below the contour curve) but sufficiently different than l (i.e. outside the band). Thus, FTAs are strong building blocs with two "smaller" and one "larger" country.

The relationship between market size asymmetry and the role of FTAs is intuitive. FTAs are strong stumbling blocs with two larger and one smaller country for two reasons. First, the world market is large enough that l does not veto global free trade in the multilateralism game where the only form of liberalization is a direct move to global free trade. Second, m and l protect valuable preferential market access as insiders, creating strong preference erosion fears and preventing global free trade in the bilateralism game. Conversely, FTAs are strong building blocs with two smaller and one larger country for opposite reasons. A small world market means l vetoes a direct move to global free trade in the multilateralism game. But, m and l protect a low degree of preferential market access as insiders meaning global free trade emerges in the bilateralism game given weak fears of preference erosion.

6 Discussion

6.1 Application to real world negotiations

By linking stylized models to real world events, one must acknowledge that real world counterexamples will surely defy the model's predictions. With this in mind, recent real world negotiations involving the US are consistent with my model. Thus, my model helps shed some light on the complex and evolving web of FTAs.

First, Proposition 5 and Figure 2 predict the pervasiveness of $g_{ml} \rightarrow g_l^H \rightarrow g^{FT}$ as the equilibrium path of FTAs leading to global free trade. This is consistent with the empirical finding of Chen and Joshi (2010, p.243-244) where two countries are more likely to form an FTA when their joint market size is larger and the larger insider is more likely to become the hub. Obvious real world examples include situations where the US is the large country and Canada is the medium country (the 1989 Canada-US FTA made them insiders) with the small country either Israel, Peru, Colombia, Jordan, Panama, Honduras or Korea.[25]

Second, the model gives an interpretation of the relationship between the order negotiations commence and the order they conclude: while the outsider begins *negotiations* with the smaller insider before the larger insider, the outsider *forms* the first FTA with the larger insider. Consider US–Canada–Colombia negotiations. Pre 2002, consistent with the equilibrium when $\delta > \hat{\delta}_{m,l}^{NE}(\alpha)$ where the largest countries remain insiders, Colombia was the outsider. However, the Colombian market oriented reforms of the 1990s and early 2000s plausibly made the Colombian market more attractive relative to the larger insider markets. Once α_{ms} and α_{ls} fall enough that $\delta < \bar{\delta}_{m,l}^{NE}(\alpha)$, the temporary hub benefits of sole reciprocal preferential access with Colombia compensate Canada for subsequent preference erosion. An interpretation is Canada beginning negotiations with Colombia, which happened in 2002.

[25] See below for details regarding Colombia and Korea. For the other countries, the US implemented FTAs with Israel, Peru, Jordan, Panama and Honduras in (respectively) 1985, 2007, 2001, 2011 and 2005 while Canada implemented FTAs with these countries in (respectively) 1997, 2009, 2012, 2013 and 2014.

Assuming $\delta > \bar{\delta}^{NE}_{l,m}(\alpha)$, the US will not initiate negotiations with Colombia if Canada does not. But, given a pre-existing US–Canada FTA, the unique equilibrium is the US becomes the hub upon anticipating a Canada–Colombia FTA. Indeed, this is consistent with history as the US initiated discussions with Colombia in 2004 that led to the 2006 US–Colombia FTA prior to the 2008 Canada–Colombia FTA. Moreover, similar interpretations apply to US, Canada, Australia and Korea negotiations.[26]

Interestingly, the model suggests an observable implication regarding FTA exclusion incentives, and the underlying fear of preference erosion, as an explanation for why FTAs do not form. Since spoke–spoke FTAs do not suffer from the fear of preference erosion that insider–outsider FTAs suffer, spoke–spoke FTAs should have a higher conditional probability of formation than insider–outsider FTAs.[27] Indeed, this observable implication receives empirical support from Chen and Joshi (2010) who find the conditional probability of a spoke–spoke FTA exceeds that of an insider–outsider FTA by a factor of four.

Finally, the discount factor δ mediates the effects of FTA exclusion incentives by affecting how much countries care about future preference erosion fears. But, what real world factors determine δ? Importantly, a period in the model is the time, say T years, needed to negotiate an agreement. Thus, denoting the one year discount factor by β, δ is really $\delta = \beta^T$. Hence, T is an important determinant of δ. An important determinant of β could be the stability of the political regime with governing parties placing more weight on future events when they are more certain they could hold power in the future. Within stable political regimes, term limits and other legislative rules shaping time in office could drive β.

6.2 Sensitivity to model assumptions

I now discuss four ways that my results are insensitive to the model's assumptions. First, my protocol is similar in spirit to Aghion et al. (2007) where a leader country (e.g. the US) makes proposals to two follower countries who cannot make proposals themselves. Indeed, by allowing the follower countries to make proposals, and thus form spoke-spoke FTAs, my protocol is more general than Aghion et al. (2007). Nevertheless, it is straightforward to see the main results are insensitive to alternative protocol orderings.

[26] For US–Canada–Korea case, note that formal Canada–Korea negotiations began in 2005 after which US–Korea negotiations began in 2006 leading the conclusion of the US–Korea FTA in 2007 before conclusion of the Canada–Korea FTA in 2014. For the US–Australia–Korea case note that the 2005 US–Australia FTA makes them insiders. Further, the 2007 US–Korea FTA lay dormant in the US Congress while Australia–Korea negotiations began in 2009 yet the US–Korea FTA then passed through Congress in 2011 before the 2014 Australia–Korea FTA.

[27] To be clear, the observable implication is $\text{pr}\left(g+jk \mid g=g_i^H\right) > \text{pr}\left(g+jk \mid g=g_{ij}\right)$. In the model, $\text{pr}\left(g+jk \mid g=g_i^H\right)=1$ with the presence of FTA exclusion incentives implying that $\text{pr}\left(g+jk \mid g=g_{ij}\right)=0$ when $\delta > \hat{\delta}^{NE}_{i,j}(\alpha)$.

Underlying the strong stumbling–strong stumbling bloc analysis is that, in the bilateralism game, global free trade does not emerge when $\delta > \hat{\delta}^{NE}_{m,l}(\alpha) \equiv \max\left\{\bar{\delta}^{NE}_{m,l}(\alpha), \bar{\delta}^{NE}_{l,m}(\alpha)\right\}$ but can emerge once $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$. The former case arises because the permanent FTA between m and l is strictly most preferred for m and, except for the possibility of remaining a permanent outsider, for l also. Thus, g_{ml} emerges unless: (i) l credibly refuses subsequent agreements as an outsider (i.e. $v_l(g_{sm}) > v_l\left(g^{FT}\right) > v_l\left(g^{H}_{m}\right)$) and (ii) l rejects any proposal in the subgame at the empty network (i.e. $g_{sm} \succ_l g$ for $g = g_{ml}, g_{sl}, g^{FT}$) anticipating s and m will form a permanent FTA. Further, global free trade fails to emerge when $\delta < \hat{\delta}^{NE}_{m,l}(\alpha)$ if and only if these same two conditions hold.

The possibility of a permanent FTA between s and m in equilibrium is unaffected by switching m and l in the protocol ordering but, regardless of $\delta \gtrless \hat{\delta}^{NE}_{m,l}(\alpha)$, does crucially depend on s being the third proposer. Note that $g_{ml} \succ_l g$ for $g = g_{sl}, g^{FT}, \emptyset$ and $g_{ml} \succ_m g$ for $g = g_{sm}, \emptyset$ when $g_{sm} \to g_{sm}$. Thus, as the third proposer, l would propose an FTA with m while m would prefer proposing an FTA with l rather than s. Thus, if s is not the third proposer, the results actually become cleaner by removing the possibility of a permanent FTA between s and m emerging in equilibrium.

Second, the assumption of at most one agreement in a period eliminates (and only eliminates) the possibility that countries could move directly to the hub–spoke network. However, this is not driving my main result which is the strong stumbling bloc result. This result arises when $\delta > \hat{\delta}^{NE}_{m,l}(\alpha)$ and, thus, m and l remain insiders forever upon forming their FTA. But moving directly to the hub–spoke network rather than remaining an insider forever is attractive for, say, m only if $v_m\left(g^{H}_{m}\right) + \frac{\delta}{1-\delta} v_m\left(g^{FT}\right) > \frac{1}{1-\delta} v_m(g_{ml})$ which reduces to $\delta < \bar{\delta}^{NE}_{m,l}(\alpha)$. Thus, m and l prefer becoming (and remaining) insiders over a direct move to the hub–spoke network if and only if $\delta > \hat{\delta}^{NE}_{m,l}(\alpha)$.

Third, Zhang et al. (2014) show the attainment of global free trade as a stochastically stable state can depend upon the special case of three countries. However, this is not true in my model. Consider four countries A, B, C, D where each country can form one FTA per period. Take the "hub–spoke" network $g^{H} \equiv (AB, AC, BD, CD)$ where the FTAs AD and BC are the only unformed FTAs. Like earlier sections, suppose each country forms its final FTA in g^{H} so that $g^{H} \to g^{FT}$. Now take an "insider–outsider" network $g^{IO} \equiv (AB, CD)$. Notice that A prefers to form the FTA with D over the permanent status quo of g^{IO} if $v_A\left(g^{H}\right) + \frac{\delta}{1-\delta} v_A\left(g^{FT}\right) > \frac{1}{1-\delta} v_A\left(g^{IO}\right)$ which reduces to the analog of the No Exclusion condition (2) presented earlier. Thus, No Exclusion conditions and FTA exclusion incentives will still drive whether global free trade eventuates in a four country model.

Fourth, unlike my multilateralism game, Saggi and Yildiz (2010) and Saggi et al. (2013) allow two country MFN agreements (two countries agree *partial* tariff cuts but extend these to the non–member). Importantly, this can undermine global free trade in the absence of FTAs, thus mitigating the role of FTAs as strong stumbling blocs, by creating incentives for free riding on the MFN tariff reductions of others. However, when (i) global free trade is the equilibrium of my multilateralism game and (ii) each country prefers global free trade over being a member of a two country MFN agreement that allows the non-member country to free ride on the MFN tariff concessions, Proposition 6 in the Supplemental Appendix shows free riding on a two country MFN agreement is not an equilibrium in my model. These conditions are satisfied for the areas of the parameter spaces identified in Lemma 1. Thus, adding two country MFN agreements does not affect my strong stumbling bloc result.[28]

7 Conclusion

This paper uses a dynamic farsighted network formation model to analyze the long standing issue of whether FTAs prevent or facilitate the attainment of global free trade, i.e. whether FTAs are building blocs or stumbling blocs. Like Saggi and Yildiz (2010, 2011), I infer the role of FTAs by comparing the equilibrium outcomes of two games: one where countries can form FTAs or move directly to global free trade and one where FTAs are not possible.

Unlike Saggi and Yildiz (2010, 2011), I find FTAs can be strong stumbling blocs meaning that global free trade is only attained in the game where FTAs are not possible. This result emerges because a pair of insider countries have an FTA exclusion incentive: the insiders want to exclude the outsider from a direct move to global free trade. Fears of preference erosion create the FTA exclusion incentive; while exchanging additional reciprocal preferential access with the outsider makes becoming the hub attractive, the would-be hub anticipates an FTA between the spokes will then erode the reciprocal preferential access enjoyed as the hub. The strong stumbling bloc result emerges under symmetry but, more generally, the FTA exclusion incentive interacts with asymmetry such that FTAs are strong stumbling blocs with two larger countries and one smaller country but FTAs are strong building blocs with two smaller countries and one larger country.

While Saggi and Yildiz (2010, 2011) cannot find my strong stumbling bloc result because insiders do not hold FTA exclusion incentives in their models, I show FTA exclusion incentives emerge in numerous trade models. Moreover, while Saggi et al. (2013) find that Customs Unions can be strong stumbling blocs to global free trade, FTAs outnumber CUs by a ratio of 9:1 which places fundamental importance on the FTA analysis.

[28]Intuitively, the last proposer in the protocol proposes FT and the other countries accept. In turn, no country accepts a proposal earlier in the protocol that allows the non-member to free ride on the two country MFN concessions embodied in the proposal.

Importantly, the model yields predictions consistent with real world FTA formation and FTA non–formation. The model provides interpretations of recent FTA negotiations involving the US by relating the path of FTAs to (i) country asymmetries (matching empirical findings of Chen and Joshi (2010)) and (ii) the order that FTA negotiations commence. In particular, commencement of negotiations between the outsider and the smaller insider induce the larger insider to become the hub. Moreover, the model suggests FTA exclusion incentives, and the underlying preference erosion fears, help explain FTA non-formation. An observable implication is the conditional probability of spoke–spoke FTAs should exceed that of insider–outsider FTAs which receives empirical support from Chen and Joshi (2010).

Finally, the model suggests ambiguities in GATT Article XXIV may actually promote global free trade by mitigating preference erosion fears. Allowing FTA members to omit some industries from an FTA and phase in tariff removal over time may increase the immediate benefit of the FTA to the extent that the hub benefits outweigh preference erosion fears.

Appendix

A Underlying trade models

Competing exporters model with market size asymmetry. The original version of the competing exporters model dates back to Bagwell and Staiger (1999). Three countries are denoted by $i = s, m, l$ and three (non-numeraire) goods are denoted by $Z = S, M, L$. Demand for good Z in country i is given by $d_i\left(p_i^Z\right) = \bar{d}_i - p_i^Z$ where p_i^Z is the price of good Z in country i. Again, country i's characteristic is $\alpha_i \equiv \bar{d}_i$. Each country i has an endowment $e_i^Z > 0$ of goods $Z \neq I$ and an endowment $e_i^Z = 0$ of good $Z = I$. I assume symmetric endowments so that $e_i^Z = e$ for all i and $Z \neq I$. Thus, country i has a "comparative disadvantage" in good I while countries j and k have a "comparative advantage" in good I and, in equilibrium, compete with each other when exporting good I to country i.

Ruling out prohibitive tariffs, the equilibrium price of any good I is linked across countries by no-arbitrage conditions: $p_i^I = p_j^I + \tau_{ij} = p_k^I + \tau_{ik}$. International market clearing conditions yield closed form solutions for equilibrium prices. Letting $x_i^Z = e_i^Z - d_i\left(p_i^Z\right)$ denote country i's *net* exports of good Z, then market clearing in good Z requires $\sum_i x_i^Z = 0$. This yields:

$$p_i^I = \frac{1}{3}\left[\sum_{h\in N} \bar{d}_h - 2e + \tau_{ij} + \tau_{ik}\right] \text{ and } p_j^I = \frac{1}{3}\left[\sum_{h\in N} \bar{d}_h - 2e + \tau_{ik} - 2\tau_{ij}\right] \text{ for } j \neq i.$$

Competing importers model with flexible supply. The competing importers model was introduced by Horn et al. (2010) and extended to a three country setting by Missios et al. (2014). Again, three countries are denoted by $i = s, m, l$ and three (non-numeraire) goods are denoted by $Z = S, M, L$ with demand identical to the competing exporters model. However, unlike the competing exporters model, the competing importers model presented here features flexible supply with the supply of good Z by country i given by $x_{ii}^Z\left(p_i^Z\right) = \lambda_i^Z p_i^Z$. Thus, $\frac{1}{\lambda_i^Z}$ represents the slope of this supply curve. More specifically, $\lambda_i^Z = 1$ for $Z \neq I$ but $\lambda_i^I = 1+\lambda_i$ where $\lambda_i > 0$. Thus, countries j and k have a "comparative disadvantage" in good I and, in equilibrium, compete for imports of good I from country i who has a "comparative advantage" in good I and, in equilibrium, is the sole exporter of good I. In this model, country i's characteristic is $\alpha_i \equiv \bar{d}_i$ under market size asymmetry and symmetric technology but $\alpha_i \equiv \frac{1}{\lambda_i^I}$ under asymmetric technology and symmetric market size.

Ruling out prohibitive tariffs, the equilibrium price of any good I is linked across countries by no-arbitrage conditions: $p_j^I = p_i^I + \tau_{ji}$ and $p_k^I = p_i^I + \tau_{ki}$. International market clearing conditions yield closed form solutions for equilibrium prices. Letting $m_{ji}^I = d_j\left(p_j^I\right) - x_{jj}^I\left(p_j^I\right)$ denote country j's imports of good I from country i and $x_{ij}^I = x_{ii}^I\left(p_i^I\right) - d_i\left(p_i^I\right) - m_{ki}^I$ denote country i's exports of good I to country j, then market clearing in good I requires $x_{ij}^I = m_{ji}^I$ and $x_{ik}^I = m_{ki}^I$. This yields:

$$p_i^I = \frac{1}{6+\lambda_i}\left(\sum_{h\in N}\bar{d}_h - 2\tau_{ji} - 2\tau_{ki}\right) \text{ and } p_j^I = \frac{1}{6+\lambda_i}\left(\sum_{h\in N}\bar{d}_h - 2\tau_{ki} + (4+\lambda_i)\,\tau_{ji}\right) \text{ for } j \neq i.$$

Let $W_i \equiv CS_i + PS_i + TR_i$ denote the national welfare of country i where CS_i, PS_i and TR_i denote country i's consumer surplus, producer surplus and tariff revenue. In the political economy oligopolistic model $v_i \equiv \pi_i$, but $v_i \equiv W_i$ in all other models.

Political economy oligopolistic model. Let $\tilde{\alpha}^2 \equiv \sum_{i\in N}\alpha_i^2$ and let the common exogenous tariff be τ. Then, $\pi_i\left(g_{ij}\right) = \frac{1}{16}\left[(\alpha_i+\tau)^2 + (\alpha_j+\tau)^2 + (\alpha_k - 2\tau)^2\right]$ which reduces to $\pi_i\left(g_{ij}\right) = \frac{1}{16}\left[\tilde{\alpha}^2 + 2\tau\left(\alpha_i+\alpha_j\right) - 4\tau\alpha_k + 6\tau^2\right]$. Similarly, $\pi_i\left(g_i^H\right) = \frac{1}{16}\left[\tilde{\alpha}^2 + 2\tau\left(\alpha_j+\alpha_k\right) + 2\tau^2\right]$, $\pi_i\left(g^{FT}\right) = \frac{1}{16}\tilde{\alpha}^2$, $\pi_i\left(\emptyset\right) = \frac{1}{16}\left[\tilde{\alpha}^2 + 4\tau\alpha_i - 4\tau\left(\alpha_j+\alpha_k\right) + 12\tau^2\right]$, $\pi_i\left(g_k^H\right) = \frac{1}{16}\left[\tilde{\alpha}^2 + 2\tau\alpha_i - 6\tau\alpha_j + 10\tau^2\right]$, $\pi_i\left(g_{jk}\right) = \frac{1}{16}\left[\tilde{\alpha}^2 + 4\tau\alpha_i - 6\tau\left(\alpha_j+\alpha_k\right) + 22\tau^2\right]$. Note, $\pi_i\left(g\right) - \pi_i\left(g'\right)$ always reduces to a simple expression like $\pi_i\left(g_{ij}\right) - \pi_i\left(g^{FT}\right) \propto \alpha_i + \alpha_j - 2\alpha_k + 3\tau$. Moreover, non-prohibitive tariffs require $\tau < \frac{\alpha_s}{3}$ since the binding constraint on $x_{ij}^*\left(g\right) > 0$ is given by $x_{ij}^*\left(g_{jk}\right) = \frac{1}{4}\left[\alpha_j - 3\tau\right] > 0$.

Oligopolistic model. For arbitrary tariffs: $CS_i = \frac{1}{32}\left(3\alpha_i - \sum\limits_{h=j,k}\tau_{ih}\right)^2$, $PS_i = \frac{1}{16}\left(\alpha_i + \sum\limits_{h=j,k}\tau_{ih}\right)^2 + \frac{1}{16}\sum\limits_{h=j,k;h'\neq h,i}\left(\alpha_h + \tau_{hh'} - 3\tau_{hi}\right)^2$ and $TR_i = \frac{1}{4}\sum\limits_{h=j,k;h'\neq h,i}\tau_{ih}\left(\alpha_i + \tau_{ih'} - 3\tau_{ih}\right)$. Country i's optimal tariffs are $\bar{\tau}_i\left(g\right) = \frac{3\alpha_i}{11\eta_i(g)-1}$ which reduces to $\bar{\tau}_i\left(\emptyset\right) = \bar{\tau}_i\left(g_{jk}\right) = \frac{3}{10}\alpha_i$ and $\bar{\tau}_i\left(g_{ij}\right) =$

$\bar{\tau}_i\left(g_j^H\right) = \frac{1}{7}\alpha_i$.

Competing exporters model. For arbitrary tariffs: $CS_i = \frac{1}{18}\left(2\alpha_i - \alpha_k - \alpha_j + 2e - \tau_{ij} - \tau_{ik}\right)^2 + \frac{1}{18}\sum_{h=j,k;h'\neq h,i}\left(2\alpha_i - \alpha_k - \alpha_j + 2e + 2\tau_{hi} - \tau_{hh'}\right)^2$, $PS_i = \frac{1}{3}e\sum_{h=j,k;h'\neq h,i}\left(\alpha_i + \alpha_j + \alpha_k - 2e - 2\tau_{hi} + \tau_{hh'}\right)$ and $TR_i = \frac{1}{3}\sum_{h=j,k;h'\neq h,i}\tau_{ih}\left(\alpha_i + \alpha_{h'} - 2\alpha_h + e - 2\tau_{ih} + \tau_{ih'}\right)$. Optimal tariffs are given by $\tau_{ik}(\emptyset) = \tau_{ik}(g_{jk}) = \frac{1}{8}\left(2\alpha_i - \alpha_j - \alpha_k + 2e\right)$ and $\tau_{ik}(g_{ij}) = \tau_{ik}\left(g_j^H\right) = \frac{1}{11}\left(\alpha_i + 4\alpha_j - 5\alpha_k + e\right)$. Making the normalization $e = 1$, non-negative tariffs require $\alpha_{ls} < 1 + \frac{1}{5}\alpha_{ms}$. Together with the requirement of $\alpha_{ls} \geq \alpha_{ms}$, this implies $\alpha_{ls} \leq \frac{5}{4}$.

Competing importers model. For arbitrary tariffs: $CS_i = \frac{1}{2}\left(\frac{1}{6+\lambda_i}\right)^2\left[(5+\lambda_i)\bar{d}_i - \bar{d}_j - \bar{d}_k\right.$ $\left.2\left(\tau_{ji} + \tau_{ki}\right)\right]^2 + \frac{1}{2}\sum_{h=j,k;h'\neq h,i}\left(\frac{1}{6+\lambda_h}\right)^2\left[(5+\lambda_h)\bar{d}_i - \bar{d}_j - \bar{d}_k + 2\tau_{h'h} - (4+\lambda_h)\tau_{ih}\right]^2$, $PS_i = \frac{1}{2}\frac{1+\lambda_i}{(6+\lambda_i)^2}\left(\bar{d}_i - 2\left(\tau_{ji} + \tau_{ki}\right)\right)^2 + \frac{1}{2}\sum_{h=j,k;h'\neq h,i}\frac{1}{(6+\lambda_h)^2}\left(\bar{d}_h - 2\tau_{h'h} + (4+\lambda_h)\tau_{ih}\right)^2$ and $TR_i = \sum_{h=j,k;h'\neq h,i}\frac{\tau_{ih}}{6+\lambda_h}\left[(4+\lambda_h)\bar{d}_i - 2\bar{d}_j - 2\bar{d}_k + 4\tau_{h'h} - 2(4+\lambda_h)\tau_{ih}\right]$. Optimal tariffs are given by $\tau_{ik}(\emptyset) = \tau_{ik}(g_{ij}) = \frac{\bar{d}_i\left(\lambda_k^2 + 10\lambda_k + 20\right) - 2\bar{d}_j(4+\lambda_k) - 2\bar{d}_k(6+\lambda_k)}{(6+\lambda_k)\left(\lambda_k^2 + 12\lambda_k + 28\right)}$ and $\tau_{ik}(g_{jk}) = \tau_{ik}\left(g_j^H\right) = \frac{\bar{d}_i(4+\lambda_k) - 2\left(\bar{d}_j + \bar{d}_k\right)}{(4+\lambda_k)(8+\lambda_k)}$. Under symmetric market size, optimal tariffs and exports are always strictly positive. Under symmetric technology, with $\lambda_i = 1$ for all i and $\bar{d}_s$ normalized to 1, non-negative exports require $\alpha_{ls} < \frac{3}{2}$.

Proof of Lemma 1

Political economy oligopolistic model. Under symmetry, it is straightforward to verify Condition 1 and Condition 3.

Under asymmetry, it is straightforward to verify $\pi_i\left(g_i^H\right) - \pi_i\left(g^{FT}\right) > 0$ as well as parts (i), (iv) and (v) of Condition 2 and that the binding constraint on parts (ii) and (iii) is $\pi_l\left(g^{FT}\right) > \pi_l\left(g_m^H\right)$ which reduces to $\alpha_l < 3\alpha_s - 5\tau$. However, one can also impose $\pi_l\left(g_m^H\right) > \pi_l(g_{sm})$ which reduces to $\alpha_l < 3\alpha_m - 6\tau$. Then, part (vi) is irrelevant and $\pi_l\left(g^{FT}\right) > \pi_l\left(g_m^H\right) > \pi_l(g_{sm})$ makes Condition 3(iii) irrelevant. Numerically, one can show Condition 3(i) holds because it holds for $\delta = \underline{\delta}$ where $\underline{\delta}$ is the minimum of 1 and $\underset{\delta}{\operatorname{argmin}}\,\pi_i(g_{ij}) + \delta\pi_i\left(g_j^H\right) + \frac{\delta^2}{1-\delta}\pi_i\left(g^{FT}\right) - \frac{1}{1-\delta}\pi_i(\emptyset)$. Condition 3(ii) holds for any δ when $\pi_l\left(g^{FT}\right) > \pi_l(\emptyset)$ but $\pi_l\left(g^{FT}\right) > \pi_l(\emptyset)$ implies some critical δ, say $\tilde{\delta}(\alpha) < 1$, such that part (ii) fails for $\delta > \tilde{\delta}(\alpha)$. Nevertheless, this never binds in equilibrium because $\tilde{\delta}(\alpha) > \hat{\delta}_{m,l}^{NE}(\alpha)$. For Condition 4, part (i) follows from $\bar{\delta}_{i,j}^{NE}(\alpha) < \bar{\delta}_{j,i}^{NE}(\alpha)$ when $\alpha_i > \alpha_j$ and $\bar{\delta}_{j,i}^{NE}(\alpha) < \bar{\delta}_{k,i}^{NE}(\alpha) < \bar{\delta}_{k,j}^{NE}(\alpha)$ when $\alpha_i > \alpha_j > \alpha_k$. Part (ii) follows given one can verify that $\bar{\delta}_{l,s}^{NE}(\alpha) > \bar{\delta}_{m,s}^{FT-O}(\alpha)$ when $\alpha_{ls} \geq \alpha_{ms}$. Finally, for part (iii), it is trivial to verify $\frac{\partial\bar{\delta}_{i,j}^{NE}(\alpha)}{\partial\alpha_{hk}} < 0$ for $h = i, j$.

Oligopolistic model with endogenous tariffs. For the purposes of part (iv) of Condition 2 and Conditions 3-4, let $\alpha_1 \equiv (\alpha_{ls}, \alpha_{ms}) = (1.35, 1.25)$ and $\alpha_2 = (1.6, 1.25)$ denote two particular parameter vectors noting that $W_l\left(g^{FT}\right) > W_l(\emptyset)$ only for α_1.

Condition 2: For the effect of FTAs on members, i.e. parts (ii)-(iii), first note that $W_i(g_{ij}) - W_i(\emptyset) \propto -1.37\alpha_i^2 + 2.29\alpha_j^2 > 0$. Thus, $\alpha_{ms} \lesssim 1.29$ and $\alpha_{ls} \lesssim 1.29\alpha_{ms} \lesssim 1.67$ imply $W_i(g_{ij}) > W_i(\emptyset)$ except potentially when $i = l$ and $g_{ij} = g_{sl}$. Hereafter, let $\alpha_{ms} \lesssim 1.29$ and $\alpha_{ls} \lesssim 1.29\alpha_{ms} \lesssim 1.67$. More generally, an FTA between i and j has the following effects on members: $W_i\left(g_i^H\right) - W_i(g_{ij}) \propto -.43\alpha_i^2 + 2.29\alpha_j^2 > 0$, $W_i\left(g_j^H\right) - W_i(g_{jk}) \propto -1.37\alpha_i^2 + 1.35\alpha_j^2 > 0$ for $\alpha_j \gtrsim 1.01\alpha_i$ and $W_i\left(g^{FT}\right) - W_i\left(g_k^H\right) \propto -.43\alpha_i^2 + 1.35\alpha_j^2 > 0$. Note, part (i) follows since $W_i(g + ij)$ is increasing in α_j. For the effects of FTAs on non-members, $W_i\left(g^{FT}\right) - W_i\left(g_i^H\right) \propto -.61\left(\alpha_j^2 + \alpha_k^2\right) < 0$. Parts (v)-(vi) follow because $W_i(g_{ij}) - W_h\left(g^{FT}\right) \propto .43\alpha_i^2 + .61\alpha_j^2 - 1.68\alpha_k^2$ which is > 0 for $g_{ij} = g_{ml}$ given α_1 or α_2 but < 0 for $g_{ij} = g_{sm}$. Part (iv) follows because $W_i\left(g^{FT}\right) - W_i(\emptyset) \propto -1.8\alpha_{ik}^2 + 1.68\left(\alpha_{jk}^2 + 1\right)$ is < 0 only if $i = l$ and is decreasing in α_{ik} and increasing in α_{jk}.

Conditions 3-4: First consider Condition 3. For any α satisfying Condition 2, part (i) holds given $W_h\left(g_l^H\right) > W_h(\emptyset)$ for $h = s, m$. Part (ii) holds for any α satisfying Condition 2 and $W_l\left(g^{FT}\right) > W_l(\emptyset)$, including α_1, but only for $\delta \lesssim .82$ given α_2 (which is not binding in equilibrium given $\hat{\delta}_{m,l}^{NE}(\alpha_2) \approx .76$). Part (iii) is only relevant for α_2 given $W_l\left(g^{FT}\right) > W_l(g_{sm})$ for α_1 and holds for all δ when $h = s, m$. Finally, consider Condition 4. Part (iii) holds because, for any α satisfying Condition 2, $\frac{\partial \bar{\delta}_{i,j}^{NE}(\alpha)}{\partial \alpha_{hk}} < 0$ for $h = i, j$. Parts (i)-(ii) follow from $\hat{\delta}_{m,l}^{NE}(\alpha) < 1$, $\bar{\delta}_{m,s}^{NE}(\alpha) > 1$ and $\bar{\delta}_{l,s}^{NE}(\alpha) > 1$ for $\alpha = \alpha_1, \alpha_2$.

Competing exporters model with market size asymmetry. Consider the parameter vector $\alpha_1 \equiv (\alpha_{ls}, \alpha_{ms}) = (1.11, 1.04)$ noting that $W_l\left(g^{FT}\right) - W_l(\emptyset) > 0$ for α_1.

Condition 2: This is straightforward to verify. In particular, only m and l have an FTA exclusion incentive. Moreover, for any α satisfying non-negative optimal tariffs, part (i) follows because $W_i\left(g_j^H\right) - W_i\left(g_k^H\right) \propto \alpha_j - \alpha_k > 0$ iff $\alpha_j > \alpha_k$ and $W_i(g_{ij}) - W_i(g_{jk}) \propto \alpha_j - \alpha_k > 0$ iff $\alpha_j > \alpha_k$ and, for the purposes of part (iv), $\frac{\partial\left[W_l\left(g^{FT}\right) - W_l(\emptyset)\right]}{\partial \alpha_{ls}} \propto 10\alpha_{ls} - 5\alpha_{ms} - 41 < 0$ and $\frac{\partial\left[W_l\left(g^{FT}\right) - W_l(\emptyset)\right]}{\partial \alpha_{ms}} \propto -5\alpha_{ls} + \alpha_{ms} + 22 > 0$.

Conditions 3-4: For α_1, parts (i)-(ii) of Condition 3 hold given Condition 2, $W_i\left(g_j^H\right) > W_i(\emptyset)$ when $\alpha_j > \alpha_i$, and $W_l\left(g^{FT}\right) > W_l(\emptyset)$. Part (iii) holds for all δ when $h = s, m$. For Condition 4, parts (i) and (ii) follow from Condition 2 and $\bar{\delta}_{i,s}^{NE}(\alpha_1) > 1$ for $i = m, l$ while part (iii) holds because $\frac{\partial \bar{\delta}_{i,j}^{NE}(\alpha_1)}{\partial \alpha_{hk}} < 0$ for $k = s$ and $h \neq k$.

Competing importers model. Under symmetry, Missios et al. (2014) have shown Condition 1. Thus, only part i) of Condition 3 needs verification. Indeed, this holds given it holds for $\delta = \underline{\delta}(\lambda, \alpha)$ where $\underline{\delta}(\lambda, \alpha) \equiv \underset{\delta}{\operatorname{argmin}}\, v_i(g_{ij}) + \delta v_i\left(g_j^H\right) + \frac{\delta^2}{1-\delta} v_i\left(g^{FT}\right) - \frac{1}{1-\delta} v_i(\emptyset)$.

For market size and technology asymmetry respectively, a parameter vector is $\alpha^d \equiv \left(\bar{d}_l, \bar{d}_m, \bar{d}_s, \lambda\right)$ and $\alpha_1^\lambda \equiv \left(\lambda_l, \lambda_m, \lambda_s, \bar{d}\right)$. Consider the parameter vectors $\alpha_1^d = (1.01, 1.005, 1, 1)$, $\alpha_2^d = (1.03, 1.005, 1, 1)$, $\alpha_1^\lambda = (.95, .96, 1, 1)$ and $\alpha_2^\lambda = (.85, .96, 1, 1)$. Note, $W_i\left(g^{FT}\right) <$

$W_i(\varnothing)$ only for α_2^d and α_2^λ and only for $i = l$.

Condition 2: This is straightforward to verify for α_1^d, α_2^d, α_1^λ and α_2^λ. Any FTA is mutually beneficial for members but imposes negative externalities on non-members. Thus, parts (ii)-(iii) hold and part (vi) is irrelevant. Part (i) holds for α_1^λ and α_2^λ and holds for any α^d satisfying non-negative optimal tariffs. Part (v) holds for α_1^d, α_2^d, α_1^λ and α_2^λ. For part (iv), $\frac{\partial\left[W_l\left(g^{FT}\right)-W_l(\varnothing)\right]}{\partial\alpha_{ls}} \propto -1894\left(1+\alpha_{ms}\right)+1212\alpha_{ls} < 0$ and $\frac{\partial\left[W_l\left(g^{FT}\right)-W_l(\varnothing)\right]}{\partial\alpha_{ms}} \propto 1792+1432\alpha_{ms} - 1894\alpha_{ls} > 0$ for any α^d satisfying non-negative optimal tariffs while $\frac{\partial\left[W_l\left(g^{FT}\right)-W_l(\varnothing)\right]}{\partial\alpha_{ls}} < 0$ and $\frac{\partial\left[W_l\left(g^{FT}\right)-W_l(\varnothing)\right]}{\partial\alpha_{ms}} > 0$ for α_1^λ and α_2^λ.

Conditions 3-4: For Condition 3, part (i) holds for any δ given α_1^d, α_2^d, α_1^λ or α_2^λ. Part (ii) holds for any δ given α_1^d or α_1^λ but part (ii) only holds for α_2^d or α_2^λ when $\delta \lesssim .89$ and $\delta \lesssim .94$ respectively (which are not binding in equilibrium given $\hat{\delta}_{m,l}^{NE}\left(\alpha_2^d\right) \approx .57$ and $\hat{\delta}_{m,l}^{NE}\left(\alpha_2^\lambda\right) \approx .6$). Part (iii) and part (ii) of Condition 4 are irrelevant because $W_i\left(g^{FT}\right) > W_i\left(g_j^H\right) > W_i\left(g_{jk}\right)$ (and thus $\bar{\delta}_{i,j}^{FT-O}(\alpha) < 0$) for any i, j, k and α_1^d, α_2^d, α_1^λ or α_2^λ. Finally, it is trivial to verify part (i) of Condition 4 and $\frac{\partial\hat{\delta}_{m,l}^{NE}}{\partial\alpha_{ls}} + \frac{\partial\hat{\delta}_{m,l}^{NE}}{\partial\alpha_{ms}} < 0$ for α_1^d, α_2^d, α_1^λ or α_2^λ.

B Proofs

Proof of Lemma 2

Given $v_h\left(g^{FT}\right) > v_h\left(g_i^H\right)$ for $h \neq i$ and $v_i\left(g_i^H\right) > v_i\left(g^{FT}\right)$ then, regardless of the position of i, j and k in the protocol, $r_i\left(g_i^H, FT\right) = N$ and $r_h\left(g_i^H, jk\right) = Y$ for $h = j, k$. Thus, $\rho_h\left(g_i^H\right) = jk$ for $h \neq i$ and $\rho_i\left(g_i^H\right) = \phi$. Therefore, $g_i^H \to g^{FT}$.■

Proof of Lemma 3

Let $\delta > \bar{\delta}^{NE}$. Then, by definition, g_{ij} is strictly most preferred for i and j. Thus, in stage 3, $\rho_h\left(g_{ij}\right) = \phi$ if $h \neq k$ and $r_h\left(g_{ij}, \rho_k\left(g_{ij}\right)\right) = N$ for $h \neq k$ and $\rho_k\left(g_{ij}\right) \in \{hk, FT\}$. In turn, the same logic applies in stage 2 and stage 1. Therefore, $g_{ij} \to g_{ij}$.

Now let $\delta < \bar{\delta}^{NE}$. Then, (i) $g_h^H \succ_h g_{ij} \succ_h g^{FT} \succ_h g_{h'}^H$ for $h \neq k$, $h' \neq k$ and $h \neq h'$ and (ii) $g^{FT} \succ_k g_h^H \succ_k g_{ij}$ for $h \neq k$. Without knowing the position of the outsider k in the protocol, there are three cases to consider. But, without loss of generality, let i be the proposer before j.

First, let the outsider k be the proposer in stage 3. In stage 3, $r_i\left(g_{ij}, FT\right) = N$ for $h = i, j$. But, $r_h\left(g_{ij}, hk\right) = Y$ for $h = i, j$ and thus $\rho_k\left(g_{ij}\right) = hk$ for some $h \neq k$. In stage 2, $\rho_j\left(g_{ij}\right) = jk$ given that $r_k\left(g_{ij}, jk\right) = Y$ and, similarly, $\rho_i\left(g_{ij}\right) = ik$ in stage 1 given that $r_k\left(g_{ij}, ik\right) = Y$. Therefore, $g_{ij} \to g_i^H$. Second, let the outsider k be the proposer in stage 2. Similar logic reveals $g_{ij} \to g_i^H$. Third, let the outsider k be the first proposer. Similar logic reveals $g_{ij} \to g_i^H$ or $g_{ij} \to g_j^H$.■

Proof of Proposition 1

For the subgame at hub-spoke networks g_i^H, Condition 1 and Lemma 2 imply $g_i^H \to g^{FT}$. Now roll back to subgames at insider-outsider networks. Lemma 3 implies $g_{ij} \to g_{ij}$ if $\delta > \bar{\delta}^{NE}$. However, given the protocol, Lemma 3 and $\delta < \bar{\delta}^{NE}$ imply $g_{ij} \to g_l^H$ if $g_{ij} = g_{hl}$ but either $g_{ij} \to g_s^H$ or $g_{ij} \to g_m^H$ if $g_{ij} = g_{sm}$.

Now roll back to the subgame at the empty network. First, let $\delta > \bar{\delta}^{NE}$. Then, Conditions 1(i) and 1(iii) imply $g_{ij} \succ_h g$ for $h = i, j$ and $g = g_{jk}, g^{FT}, \emptyset$. Thus, due to symmetry, $\rho_s(\emptyset) = sh$ for some $h = m, l$ in stage 3 given that $r_h(\emptyset, hs) = Y$ for $h = m, l$. Similar logic applies in stage 2 and stage 1 and therefore, due to symmetry, $\emptyset \to g_{sl}$ or $\emptyset \to g_{ml}$. Thus, the equilibrium path of networks is g_{ml} or g_{sl}.

Second, let $\delta \in \left(\bar{\delta}, \bar{\delta}^{NE}\right)$. Then, given Conditions 1 and 3, (i) $g_{hl} \succ_l g^{FT} \succ_l g$ for $h \neq l$ and $g = g_{sm}, \emptyset$, (ii) $g_{sm} \succ_m g^{FT} \succ_m g$ for $g = g_{ml}, g_{sl}, \emptyset$, and (iii) g^{FT} is strictly most preferred for s. In stage 3, $\rho_s(\emptyset) = FT$ given that $r_h(\emptyset, FT) = Y$ for $h \neq s$. In stage 2, given the FT outcome in stage 3, $r_s(\emptyset, sm) = N$ but $r_h(\emptyset, FT) = Y$ for $h = s, l$. Thus, $\rho_m(\emptyset) = FT$. In stage 1, given the FT outcome in stage 2, $r_h(\emptyset, hl) = N$ for $h \neq l$ but $r_h(\emptyset, FT) = Y$ for $h = s, m$. Thus, $\rho_l(\emptyset) = FT$. Therefore, $\emptyset \to g^{FT}$. Thus, g^{FT} is the equilibrium path of networks.

Finally, let $\delta < \bar{\delta}$. This leaves l's preferences unchanged relative to $\delta \in \left(\bar{\delta}, \bar{\delta}^{NE}\right)$ but now (i) $g_{sm} \succ_m g_{ml} \succ_m g^{FT} \succ_m g$ for $g = g_{sl}, \emptyset$, and (ii) $g_{sh} \succ_s g^{FT} \succ_s g$ for $h \neq s$ and $g = g_{ml}, \emptyset$. In stage 3, $r_h(\emptyset, sh) = Y$ for $h \neq s$ and thus $\rho_s(\emptyset) = sl$ or $\rho_s(\emptyset) = sm$. In stage 2, $\rho_m(\emptyset) = sm$ given that $r_s(\emptyset, sm) = Y$. In stage 1, the outcome of sm in stage 2 implies $r_m(\emptyset, ml) = N$ but $r_s(\emptyset, sl) = Y$. Thus, $\rho_l(\emptyset) = sl$ and therefore $\emptyset \to g_{sl}$. Hence, the equilibrium path of networks is $g_{sl} \to g_l^H \to g^{FT}$.■

Proof of Proposition 2

Let $v_i\left(g^{FT}\right) > v_i(\emptyset)$ for all i. In stage 3, $r_i(\emptyset, FT) = Y$ for $i \neq s$ and thus $\rho_s(\emptyset) = FT$. Similar logic applies in stage 2 and stage 1. Therefore $\emptyset \to g^{FT}$ and the equilibrium path of networks is g^{FT}. Now let $v_i\left(g^{FT}\right) < v_i(\emptyset)$ for some i. In stage 3, either $r_i(\emptyset, FT) = N$ for some $i \neq s$ or $\rho_s(\emptyset) = \phi$. Similar logic applies again in stage 2 and stage 1. Therefore $\emptyset \to \emptyset$ and the equilibrium path of networks is $\emptyset$.■

Proof of Proposition 3

In subgames at hub-spoke networks g_i^H, Condition 2 and Lemma 2 imply $g_i^H \to g^{FT}$. In subgames at insider-outsider networks g_{ij}, the logic of Lemma 3 implies $g_{ij} \to g_{ij}$ if $\delta > \hat{\delta}_{i,j}^{NE}$. Thus, $g_{ml} \to g_{ml}$ given $\delta > \hat{\delta}_{m,l}^{NE}$. Now roll back to the subgame at the empty network.

Regardless of the network paths emanating from subgames at g_{sm} and g_{sl}, Condition 2 implies (i) $g_{ml} \succ_m g$ for $g = g_{sl}, g_{sm}, g^{FT}, \emptyset$ and (ii) $g_{ml} \succ_l g$ for $g = g_{sl}, g^{FT}, \emptyset$. Conditions 2-3 also imply (iii) $g \succ_s \emptyset$ for $g = g_{sl}, g_{sm}, g^{FT}$. Thus, regardless of the outcome in stage 3,

$\rho_m(\emptyset) = ml$ in stage 2 iff $r_l(\emptyset, ml) = Y$ noting that $r_l(\emptyset, ml) = Y$ if $g_{ml} \succ_l g_{sm}$. Hence, let $g_{ml} \succ_l g_{sm}$. Then, in stage 1, $\rho_l(\emptyset) = ml$ given that $r_m(\emptyset, ml) = Y$. Thus $\emptyset \to g_{ml}$ and, hence, g_{ml} is the equilibrium path of networks.

Now let $g_{sm} \succ_l g_{ml}$. Together with Condition 2, $g_{sm} \to g_{sm}$ in the subgame at g_{sm} and $v_l(g_{sm}) > v_l(g^{FT}) > v_l(g_m^H)$. In turn, Condition 2 implies $v_s(g^{FT}) > v_s(g_{sm})$ and hence, given Condition 3, $g \succ_s g_{sm}$ for $g = g_{sl}, g^{FT}$. Thus, in stage 1, $\rho_s(\emptyset) = sm$ iff $r_l(\emptyset, sl) = r_l(\emptyset, FT) = N$ and, hence, $\emptyset \succ_l g$ for $g = g_{sl}, g^{FT}$. Letting $g \succ_l \emptyset$ for some $g = g_{sl}, g^{FT}$ then $r_l(\emptyset, ml) = Y$ and $\rho_m(\emptyset) = ml$ in stage 2 and $r_m(\emptyset, ml) = Y$ and $\rho_l(\emptyset) = ml$ in stage 1. Thus, $\emptyset \to g_{ml}$ and g_{ml} is the equilibrium path of networks. Conversely, let $\emptyset \succ_l g$ for $g = g_{sl}, g^{FT}$. Then, (i) $\rho_m(\emptyset) = sm$ in stage 2 given $r_l(\emptyset, ml) = r_l(\emptyset, FT) = N$ but $r_s(\emptyset, sm) = Y$ and (ii) and $\rho_l(\emptyset) = \phi$ in stage 1. Thus, $\emptyset \to g_{sm}$ and g_{sm} is the equilibrium path of networks.■

Proof of Proposition 4

In subgames at hub-spoke networks g_i^H, Condition 2 and Lemma 2 imply $g_i^H \to g^{FT}$. In turn, the only way that g^{FT} does not eventually emerge from some insider-outsider network g_{ij} is if $g_{ij} \to g_{ij}$. The next part of the proof shows that, once $\delta < \hat{\delta}_{m,l}^{NE}(\alpha)$, this is only possible for the subgame at $g_{ij} = g_{sm}$.

First, consider $g_{ij} = g_{ml}$ noting that (i) $g^{FT} \succ_s g_l^H \succ_s g_m^H \succ_s g_{ml}$, (ii) $g_h^H \succ_h g_{ml}$ for some $h = m, l$ and (iii) $g \succ_h g^{FT} \succ_h g_{h'}^H$ for $g = g_{ml}, g_h^H$ and $h \neq s$, $h' \neq h$. In stage 3, $r_h(g_{ml}, FT) = N$ for $h = m, l$ but $r_h(g_{ml}, sh) = Y$ for some $h = m, l$. Thus, $\rho_s(g_{ml}) = sh$ for some $h = m, l$. Noting that $r_l(g_{ml}.FT) = N$ or $\rho_m(g_{ml}) \neq FT$ in stage 2 then, regardless of the equilibrium outcome in stage 2, $r_s(g_{ml}, sl) = Y$ and $\rho_l(g_{ml}) = sl$ in stage 1. Therefore, $g_{ml} \to g_l^H$.

Second, consider $g_{ij} = g_{sl}$ noting that (i) $g^{FT} \succ_m g_l^H \succ_m g$ for $g = g_{sl}, g_s^H$, (ii) $g_h^H \succ_h g_{sl}$ for some $h = s, l$ by Condition 4(i), and (iii) $g_l^H \succ_l g^{FT} \succ_l g_s^H$. Let FT be the outcome in stage 3. Then, $r_h(g_{sl}, FT) = Y$ for $h = s, l$ in stage 2 and thus $\rho_m(g_{sl}) = FT$. In turn, in stage 1, $r_h(g_{sl}, FT) = Y$ for $h = s, m$ but $r_m(g_{sl}, \rho_l(g_{sl})) = N$ for $\rho_l(g_{sl}) = ml$ and hence $\rho_l(g_{sl}) = FT$. Now let ϕ or sm be the outcome in stage 3. In turn, in stage 2, $r_h(g_{sl}, FT) = N$ for some $h = s, l$ but $r_l(g_{sl}, ml) = Y$ either because $\delta < \bar{\delta}_{m,s}^{FT-O}(\alpha)$ and, by Condition 4, $\delta < \bar{\delta}_{l,s}^{NE}(\alpha)$ or because the outcome in stage 1 is sm. Thus, $\rho_m(g_{sl}) = ml$. In turn, $\rho_l(g_{sl}) = ml$ in stage 1. Thus, either $g_{sl} \to g^{FT}$ or $g_{sl} \to g_l^H$.

Third, consider $g_{ij} = g_{sm}$ noting that (i) $g_m^H \succ_m g_{ml}$ by Condition 4(i), and (ii) $v_h(g^{FT}) > v_h(g_{sm})$ for $h = s, m$ if $v_l(g_m^H) > v_l(g_{sm})$ by Condition 2(vi). If $v_l(g_{sm}) > v_l(g^{FT})$ then $g_{sm} \succ_l g$ for $g = g_m^H, g_s^H, g^{FT}$ and hence $r_l(g_{sm}, \rho_h(g_{sm})) = N$ for $\rho_h(g_{sm}) \neq \phi$ in stages 3 and 2. In turn, $\rho_l(g_{sm}) = \phi$ in stage 1 and, therefore, $g_{sm} \to g_{sm}$. If $v_l(g^{FT}) > v_l(g_{sm}) > v_l(g_m^H)$ then Condition 2(vi) says $v_h(g^{FT}) > v_h(g_{sm})$ for $h = s, m$. Hence, $r_h(g_{sm}, FT) = Y$ for

$h \neq s$ in stage 3 and thus $g_{sm} \to g$ for some $g \neq g_{sm}$. If $v_l\left(g^{FT}\right) > v_l\left(g_m^H\right) > v_l\left(g_{sm}\right)$ then $g_{sm} \to g$ for some $g \neq g_{sm}$ either by the logic of the previous sentence or similar logic to the case for $g_{ij} = g_{ml}$. Thus, $g_{sm} \to g_{sm}$ iff $v_l\left(g_{sm}\right) > v_l\left(g^{FT}\right)$ and this is the only case where global free trade does not eventually emerge from a subgame at an insider-outsider network.

Now roll back to the subgame at the empty network noting that $g \succ_s \emptyset$ for $g = g_{sl}, g_{sm}, g^{FT}$ and $g_{sm} \succ_m \emptyset$. Given Condition 3 says $g_{ml} \succ_h \emptyset$ for $h = m, l$ then $\emptyset \to g$ for some $g \neq \emptyset$ because there is some proposal $\rho_m\left(\emptyset\right) \neq \phi$ such that $r_h\left(\emptyset, \rho_m\left(\emptyset\right)\right) = Y$ for all recipients h in stage 2. Thus, global free trade emerges eventually unless $\emptyset \to g_{sm} \to g_{sm}$. Hence, for the remainder of the proof, let $g_{sm} \to g_{sm}$ in the subgame at g_{sm} noting that this implies $g_{ml} \succ_m g_{sm}$.

Suppose $\emptyset \succ_l g$ for $g = g_{sl}, g^{FT}$. Then, in stage 3, $r_l\left(\emptyset, FT\right) = r_l\left(\emptyset, sl\right) = N$ but $r_m\left(\emptyset, sm\right) = Y$ and hence $\rho_s\left(\emptyset\right) = sm$. In stage 2, $r_s\left(\emptyset, sm\right) = Y$ but $r_l\left(\emptyset, FT\right) = N$ and $r_l\left(\emptyset, ml\right) = Y$ iff $g_{ml} \succ_l g_{sm}$. Thus, $\rho_m\left(\emptyset\right) = ml$ iff $g_{ml} \succ_l g_{sm}$ but $\rho_m\left(\emptyset\right) = sm$ otherwise. Thus, in stage 1, $r_m\left(\emptyset, ml\right) = Y$ and $\rho_l\left(\emptyset\right) = ml$ iff $g_{ml} \succ_l g_{sm}$ but $\rho_l\left(\emptyset\right) = \phi$ otherwise. Therefore, the equilibrium path of networks is $g_{ml} \to g_l^H \to g^{FT}$ if $g_{ml} \succ_l g_{sm}$ but g_{sm} otherwise. Conversely, now suppose $g \succ_l \emptyset$ for some $g = g_{sl}, g^{FT}$. Then, Conditions 2(vi) and 3(iii) say $g \succ_s g_{sm}$ for $g = g_{sl}, g^{FT}$. Hence, $r_l\left(\emptyset, \rho_s\left(\emptyset\right)\right) = Y$ for some $\rho_s\left(\emptyset\right) \in \{sl, FT\}$ and thus $\rho_s\left(\emptyset\right) = sl$ or $\rho_s\left(\emptyset\right) = FT$. Moreover, in stage 2, $r_s\left(\emptyset, sm\right) = N$ and thus $\rho_m\left(\emptyset\right) \neq sm$. Therefore, regardless of the outcome in stage 1, $\emptyset \to g$ for some $g \neq \emptyset, g_{sm}$ and global free trade eventually emerges.$\blacksquare$

Proof of Proposition 5

Proposition 3 implies global free trade can only emerge once $\delta < \hat{\delta}_{m,l}^{NE}\left(\alpha\right)$. Thus, let $\delta < \hat{\delta}_{m,l}^{NE}\left(\alpha\right)$ for the remainder of the proof. Further, the proof of Proposition 4 established that (i) $g_i^H \to g^{FT}$ for any g_i^H, (ii) $g_{ml} \to g_l^H$, (iii) $g_{sl} \to g^{FT}$ or $g_{sl} \to g_l^H$ and (iv) $g_{sm} \to g_{sm}$ iff $v_l\left(g_{sm}\right) > v_l\left(g^{FT}\right)$. Thus, (i) $g_{ml} \succ_l g$ for $g = g_{sl}, \emptyset$, (ii) $g \succ_s \emptyset$ for $g = g_{sl}, g_{sm}, g^{FT}$ and (iii) $g \succ_m \emptyset$ for $g = g_{ml}, g_{sm}, g^{FT}$. Moreover, these preferences imply $\rho_s\left(\emptyset\right) \neq \phi$ in stage 3 and thus some agreement is the outcome in stage 3. Now consider the three cases of the proposition.

(i) Note that $g \succ_s g_{sm}$ for some $g = g_{sl}, g^{FT}$. To see this, first let $g_{sm} \succ_s g_{sl}$. In this case, $g_{sm} \to g$ for some $g \neq g_{sm}$ because $g_{sm} \to g_{sm}$ iff $v_l\left(g_{sm}\right) > v_l\left(g^{FT}\right)$ which, by Condition 3(iii), implies $g_{sl} \succ_s g_{sm}$. Further, Condition 2 and $g_{sm} \to g_m^H$ implies $g_{sl} \succ_s g_{sm}$. Thus, $g_{sm} \succ_s g_{sl}$ implies $g_{sm} \to g^{FT}$. But, this requires $v_i\left(g^{FT}\right) > v_i\left(g_{sm}\right)$ for all i which implies $g^{FT} \succ_s g_{sm}$. Second, let $g_{sm} \succ_s g^{FT}$. $g_{sm} \succ_s g^{FT}$ can only hold if $g_{sm} \to g_m^H$ because (i) $g_{sm} \to g^{FT}$ implies $v_s\left(g^{FT}\right) > v_s\left(g_{sm}\right)$ and (ii) $g_{sm} \to g_{sm}$ implies $v_l\left(g_{sm}\right) > v_l\left(g^{FT}\right)$ which, by Condition 2(vi), implies $v_s\left(g^{FT}\right) > v_s\left(g_{sm}\right)$. But $g_{sm} \to g_m^H$ and Condition 2 imply $g_{sl} \succ_s g_{sm}$. Therefore, $g \succ_s g_{sm}$ for some $g = g_{sl}, g^{FT}$ and, thus, $\rho_s\left(\phi\right) = sl$ or $\rho_s\left(\phi\right) = FT$.

Let $\rho_s(\emptyset) = sl$. Then $g_{sl} \succ_s g^{FT}$ which requires $g_{sl} \to g_l^H \to g^{FT}$. In turn, Condition 2 implies $g_{ml} \succ_m g_{sl}$. Thus, in stage 2, $r_s(\emptyset, FT) = N$ but $r_l(\emptyset, ml) = Y$ and hence $\rho_m(\emptyset) = ml$. In turn, in stage 1, $r_m(\emptyset, ml) = Y$ and $\rho_l(\emptyset) = ml$. Therefore, $g_{ml} \to g_l^H \to g^{FT}$ is the equilibrium path of networks. Now let $\rho_s(\emptyset) = FT$. Since this implies $g^{FT} \succ_s g_{sm}$, then $r_s(\emptyset, sm) = N$ in stage 2 but $r_h(\emptyset, FT) = r_l(\emptyset, ml) = Y$ for $h = s, l$ and hence $\rho_m(\emptyset) = FT$ if $g^{FT} \succ_m g_{ml}$ which reduces to $\delta > \bar{\delta}_m^{l-FT}(\alpha)$ but $\rho_m(\emptyset) = ml$ if $g_{ml} \succ_m g^{FT}$ which reduces to $\delta < \bar{\delta}_m^{l-FT}(\alpha)$. In turn, in stage 1, $\rho_l(\emptyset) = FT$ and the equilibrium path of networks is g^{FT} if $g^{FT} \succ_m g_{ml}$ but $\rho_m(\emptyset) = ml$ and the equilibrium path of networks is $g_{ml} \to g_l^H \to g^{FT}$ if $g_{ml} \succ_m g^{FT}$.

(ii) By construction, $\rho_s(\emptyset) = sm$ or $\rho_s(\emptyset) = FT$ in stage 3. Let $\rho_s(\emptyset) = FT$, i.e. $g^{FT} \succ_s g_{sm}$. Then, as in case (i), the equilibrium path of networks is g^{FT} if $\delta > \bar{\delta}_m^{l-FT}(\alpha)$ but $g_{ml} \to g_l^H \to g^{FT}$ if $\delta < \bar{\delta}_m^{l-FT}(\alpha)$. Now let $\rho_s(\emptyset) = sm$ so that $g_{sm} \succ_s g^{FT}$. Also suppose that $g_{ml} \succ_m g_{sm}$. In stage 2, $r_s(\emptyset, FT) = N$ given $\rho_s(\emptyset) = sm$ in stage 3. But, given the logic in case (i), $g_{sm} \succ_s g^{FT}$ implies $g_{sm} \to g_m^H$ which requires $v_l(g_{ml}) > v_l(g^{FT}) > v_l(g_{sm})$ and, in turn, implies $g_{ml} \succ_l g_{sm}$. Thus, in stage 2, $r_l(\emptyset, ml) = Y$ and $\rho_m(\emptyset) = ml$. In turn, $r_m(\emptyset, ml) = Y$ and $\rho_l(\emptyset) = ml$ in stage 1 and $g_{ml} \to g_l^H \to g^{FT}$ is the equilibrium path of networks.

(iii) By construction, $\rho_s(\emptyset) = sm$ in stage 3. If $g_{sm} \succ_l g_{ml}$ then $v_l(g_{sm}) > v_l(g^{FT})$ and hence $r_l(\emptyset, \rho_m(\emptyset)) = N$ for $\rho_m(\emptyset) \in \{FT, ml\}$ in stage 2. In turn, $\rho_m(\emptyset) = sm$ given $r_s(\emptyset, sm) = Y$. In turn, in stage 1, $\rho_l(\emptyset) = \phi$ and g_{sm} is the equilibrium path of networks. If $g_{ml} \succ_l g_{sm}$, then $r_l(\emptyset, ml) = Y$ in stage 2. Thus, $\rho_m(\emptyset) = ml$ if $g_{ml} \succ_m g$ for $g = g_{sm}, g^{FT}$. In turn, in stage 1, $r_m(\emptyset, ml) = Y$ and $\rho_l(\emptyset) = ml$ and therefore $g_{ml} \to g_l^H \to g^{FT}$ is the equilibrium path of networks.■

I now consider a variant on the multilateralism game that allows for two country MFN agreements. Suppose each proposer country's action space in the multilateralism game is $P_i(\emptyset) = \{\phi, FT, ij^M, ik^M\}$ where, for example, ij^M indicates i announces a two country MFN agreement with j that results in the network g_{ij}^M.

Proposition 6. *Suppose (i) $v_i(g^{FT}) > v_i(\emptyset)$ for all i and (ii) $v_i(g^{FT}) > v_i(g_{ij}^M)$* for any *$i, j$. Then, g^{FT} is the equilibrium path of networks.*

Proof. In stage 3, $r_h(\emptyset, FT) = Y$ for $h = m, l$. Thus, $\rho_s(\emptyset) = FT$. In stage 2, $r_h(\emptyset, FT) = Y$ for $h = s, l$ and thus $\rho_m(\emptyset) = FT$. In stage 1, $r_h(\emptyset, FT) = Y$ for $h = s, m$ and thus $\rho_l(\emptyset) = FT$. Therefore, the equilibrium path of networks is g^{FT}. □

References

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

Bagwell, K., Staiger, R., 1999. Regionalism and multilateral tariff cooperation. In: Pigott, J., Woodland, A. (Eds.), International trade policy and the Pacific rim. Macmillan.

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

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

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

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.

Hoekman, B., 2006. Preference erosion and the Doha development agenda. In: Paugam, J.-M., Novel, A.-S. (Eds.), Reviving the special and differential treatment of developing countries in international trade. Ifri.

Horn, H., Maggi, G., Staiger, R., 2010. Trade agreements as endogenously incomplete contracts. The American Economic Review 100 (1), 394–419.

Jackson, M., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71 (1), 44–74.

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.

Levy, P., 1997. A political-economic analysis of free-trade agreements. The American Economic Review 87 (4), 506–519.

Limão, N., Olarreaga, M., 2006. Trade preferences to small developing countries and the welfare costs of lost multilateral liberalization. World Bank Economic Review 20 (2), 217–240.

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

Missios, P., Saggi, K., Yildiz, H., 2014. External trade diversion, exclusion incentives and the nature of preferential trade agreements. Mimeo.

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

Odell, J. S., 2006. Negotiating trade: Developing countries in the WTO and NAFTA. Cambridge University Press.

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

Ornelas, E., 2005b. Trade creating free trade areas and the undermining of multilateralism. European Economic Review 49 (7), 1717–1735.

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.

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

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.

Zhang, J., Cui, Z., Zu, L., 2014. The evolution of free trade networks. Journal of Economic Dynamics and Control 38, 72–86.