Journal of International Economics 180 (2017) 117–136

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

# Are global trade negotiations behind a fragmented world of "gated globalization"?☆

James Lake, Santanu Roy*

*Department of Economics, Southern Methodist University, Dallas, TX 75275-0496*

ARTICLE INFO

*Article history:*
Received 24 November 2015
Received in revised form 22 May 2017
Accepted 9 June 2017
Available online 15 June 2017

*JEL classification:*
C73
F12
F13

*Keywords:*
Free Trade Agreement
Global free trade
Multilateralism
Tariff complementarity
Binding overhang

ABSTRACT

We show that global trade negotiations can prevent global free trade. In a simple model where global tariff negotiations precede sequential Free Trade Agreements (FTAs), we show FTA formation can expand all the way to global free trade in the absence of global tariff negotiations but global free trade never emerges when global tariff negotiations precede FTA formation. This result arises precisely because global tariff negotiations successfully elicit concessions from negotiating countries. Moreover, global tariff negotiations can produce a fragmented world of "gated globalization" where some countries form FTAs that eliminate tariff barriers among themselves while outsiders continue facing higher tariffs.



## 1. Introduction

Two rules have profoundly shaped the evolution of global tariffs since the creation of the 1948 General Agreement on Tariffs and Trade (GATT). First, the Most Favored Nation (MFN) Principle of GATT Article I outlaws discrimination among trading partners by dictating a country impose the same tariff on all trading partners. Second, GATT Article XXIV provides an escape clause from the MFN principle whereby groups of countries can form a Free Trade Agreement (FTA) and only reduce tariffs on each other if members (i) eliminate their bilateral tariffs and (ii) do not raise tariffs on non-members. Interestingly, the relative importance of these two rules in driving global tariff liberalization has varied over time.

After the Uruguay Round of global tariff negotiations in 1994, the MFN principle combined with country-by-country commitments to keep tariffs below specified tariff ceilings (i.e. tariff bindings) had generated significant tariff liberalization. Indeed, at that time, the various rounds of global tariff negotiations stood as the dominant form of global tariff liberalization with FTAs relatively few and far between. Subsequently, the post-Uruguay Round world has seen an unprecedented surge of FTAs with FTAs becoming the dominant form of global tariff liberalization. Indeed, given *de facto* global free trade arises if all country pairs are linked by FTAs, FTA expansion under Article XXIV has created new hope in an alternative route to global free trade.

This changing face of global tariff liberalization has also created interest in understanding the various political and economic factors that potentially affect the incentives for FTA expansion. Given the rapid proliferation of FTAs took place after the successful 1994 Uruguay Round of global negotiations, it is important to understand how prior global negotiations influence the incentives for subsequent FTA formation under GATT Article XXIV, and how the shadow of future FTA formation may, in turn, influence the initial

☆ We would like to thank the Editor, Giovanni Maggi, and two anonymous referees for very useful and insightful comments. We would also like to thank Mostafa Beshkar, Kristy Buzzard, Richard Chisik, Klaus Desmet, Rod Ludema, James Hartigan, Emanuel Ornelas, Peri da Silva, Murat Yildiz and Maurizio Zanardi as well as seminar and conference participants at Baylor University, Indiana University, Instituto Tecnológico Autónomo de México, Kansas State University, Pontifica Universidad Javeriana, Ryerson University, Sao Paulo School of Economics, University of Oklahoma, Spring 2014 Midwest Trade Meetings, 2014 Texas Theory Camp, 2014 Annual Conference on Economic Growth and Development and 2015 UECE Lisbon Meetings in Game Theory and Applications for useful comments and discussion.

* Corresponding author.
*E-mail addresses:* jlake@smu.edu (J. Lake), sroy@smu.edu (S. Roy).

http://dx.doi.org/10.1016/j.jinteco.2017.06.003
0022-1996/

outcome of global negotiations. How would the extent of FTA formation observed today differ if the Uruguay Round had not taken place? That is, do commitments to tariff bindings during prior global negotiations help or hinder the possibility that FTA proliferation proceeds all the way to global free trade? Could global negotiations actually be the cause of what *The Economist* recently referred to as a fragmented world of "gated globalization" where FTA expansion stops far short of global free trade?[1] These are the questions addressed in this paper.

We consider a world of three symmetric countries. For our underlying trade model, we adapt the competing exporters framework of Bagwell and Staiger (1999b) to include an import competing sector and politically motivated governments. This framework has three goods with each country exporting two comparative advantage goods and importing one comparative disadvantage good. Politically motivated governments care about national welfare but place an additional weight placed on profits of the import competing sector.

To analyze the effect of global tariff negotiations (i.e. "multilateralism") on FTA formation (i.e. "regionalism"), we compare the outcomes of two extensive form games that differ only because of the presence or absence of an initial round of global tariff negotiations. In the first game, global negotiations over tariff *bindings* are followed by FTA negotiations.[2] In the second game, there are no global negotiations preceding FTA negotiations. Once FTA negotiations conclude in either game, countries choose their applied tariffs that, in turn, generate patterns of consumption and trade. Our protocol for FTA negotiations is one of sequential bilateral FTA formation according to a randomly chosen order; the protocol ensures that after any FTA is formed, all pairs of countries that have not yet formed an FTA have the option to do so. To be clear, governments are forward looking: when undertaking global tariff negotiations they anticipate the possibility of FTA formation even though they do not yet know the precise sequential order in which country pairs will form FTAs.

Our main result is that, when political economy motivations are not too strong, multilateralism prevents global free trade.[3] When global tariff negotiations precede FTA negotiations, a tariff ridden world emerges with globally negotiated tariff bindings above zero and no more than one pair of countries linked by an FTA. However, in the *absence* of global tariff negotiations, FTA formation continues until all pairs of countries are linked by FTAs and, thus, global free trade is attained. Further, when global negotiations precede FTA formation and political economy objectives are not too strong, a world of "gated globalization" emerges where members of the single FTA practice free trade between themselves but tariff barriers remain between these FTA "insiders" and the non-member "outsider" country.

The driving force behind our main result is the different level of tariff concessions given by the eventual outsider in the presence and absence of global tariff negotiations. In the absence of global tariff negotiations, the outsider has no pre-existing tariff bindings. To gain tariff concessions from the outsider, this creates incentives for insiders to form subsequent FTAs with the outsider. Thus, as long as government political economy motivations are not too strong, sequential FTA formation leads to global free trade. However, global tariff negotiations mean all countries (including the eventual outsider) pre-commit to significant tariff concessions (via tariff bindings) before FTA negotiations. These tariff concessions obtained through forward looking global negotiations are deep enough that, upon FTA formation, the insiders have no incentive to engage in subsequent FTA formation with the outsider and, thus, global free trade does not emerge. In this sense, the success of global tariff negotiations in lowering tariffs drives our result that multilateralism prevents global free trade.

In our framework, the prospect of future FTA formation creates a "shadow of future regionalism" that affects the outcome of prior global negotiations. In particular, countries negotiate lower global tariff bindings than they would if the shadow of regionalism was not looming over global negotiations. This is driven by a *multilateral* tariff complementarity effect, as in Ornelas (2008), whereby the global tariff binding that maximizes the joint payoff of all governments falls upon FTA formation. Importantly, this differs from the usual notion of *individual* tariff complementarity where FTA members reduce tariffs on non-members due to, among other things, weaker terms of trade motivations upon FTA formation.[4] When anticipating FTA formation, global tariff negotiations aggregate the incentives of potential insiders and outsiders implying that terms of trade considerations bear no imprint on global tariff bindings. Thus, multilateral tariff complementarity reflects the forces other than terms of trade motivations that drive individual tariff complementarity.

The dependence of globally negotiated tariff bindings on subsequent FTA negotiations has significant practical implications. First, the equilibrium emergence of binding overhang and individual tariff complementarity depend on the strength of political economy motivations. When such motivations are not too strong, globally negotiated tariff bindings bind the applied tariffs of FTA members and non-members, generating zero "binding overhang". Indeed, there is a range of political economy motivations where this result emerges only because governments anticipate subsequent FTA formation. Thus, farsighted global tariff negotiations preceding FTA negotiations may help explain why essentially zero binding overhang is observed in the major countries involved in the 1994 Uruguay Round such as the US, the EU and Japan. Second, in this zero binding overhang case, our model predicts that FTA members do not lower their tariff on non-members; the usual tariff complementarity effect upon FTA formation is not observed on the equilibrium path. The reason is that farsighted global tariff negotiations already incorporate any tariff complementarity effect into applied tariffs prior to FTA negotiations taking place. Third, this logic implies the interpretation of changes in trade flows upon FTA formation is complicated because the effect that FTAs have on applied tariffs may already be embedded in multilateral tariff bindings negotiated prior to FTA formation. This is especially important given policy makers actually rely on observed trade flow changes upon FTA formation to infer the welfare effects of FTAs.[5]

While our baseline analysis employs a stylized environment, Section 6 demonstrates the robustness of our main results and provides additional insights. Departing from our symmetric protocol governing FTA negotiations, Section 6.1.1 demonstrates our main results hold when a particular country pair has a higher probability than other country pairs of having the first FTA formation opportunity. By allowing an individual country to back out of global tariff negotiations and *instead* precipitate the FTA formation process without any tariff bindings, Section 6.1.2 shows how some countries can extract larger concessions *during* global negotiations. Section 6.2 shows our results are robust to imposing exogenous, rather than endogenous, tariff bindings. By *removing* the constraints of Article XXIV and allowing positive internal tariffs among FTA members,

[1] *The Economist*, Special Report, October 2013: http://www.economist.com/news/special-report/21587384-forward-march-globalisation-has-paused-financial-crisis-giving-way.

[2] In practice, global tariff negotiations are negotiations over tariff bindings rather than the actual tariffs (i.e. applied tariffs) countries set. We model global tariff negotiations in this way.

[3] The empirical protection for sale literature (e.g. Goldberg and Maggi, 1999 and Gawande and Bandyopadhyay, 2000) finds that political economy motivations of governments tend to be weak and this is the setting in which our main result applies.

[4] The phenomenon of tariff complementarity is well known in the literature (see, for example, Richardson, 1993, Bagwell and Staiger, 1999b and Ornelas, 2005b).

[5] See Bergstrand et al. (2014, p. 3).

Section 6.3 shows the degree of FTA formation and global tariff liberalization could both rise. Interestingly, Section 6.3 also shows how governments may strategically set tariff bindings so that zero internal tariffs emerge endogenously and all equilibrium outcomes are identical to that in the presence of Article XXIV constraints. Finally, Section 6.4 discusses why incorporating political motivations stemming from both import-competing and export sectors should not affect our main results.

The paper proceeds as follows. After Section 2 discusses related literature, Section 3 presents our modified version of the Bagwell and Staiger (1999b) competing exporters model. Section 3.2 describes our game theoretic approach to modeling multilateralism and regionalism. Section 4 establishes that global tariff negotiations prevent global free trade. Section 5 establishes that global tariff negotiations can produce a fragmented world of gated globalization and characterizes the tariffs that result from global tariff negotiations. Section 6 investigates the robustness of our baseline analysis using numerous extensions. Finally, Section 7 concludes. The Appendix collects all proofs.

## 2. Related literature

A large extant literature investigates how FTAs impact global tariffs involving non-members (via global negotiations or voluntary tariff concessions by FTA members) and is often couched in the terminology of how "regionalism" affects "multilateralism" or whether FTAs are "building blocs" or "stumbling blocs" (Bhagwati, 1991, 1993) to global free trade.[6] In contrast, we ask how "multilateralism" affects "regionalism"; in particular, we ask whether multilateralism is a building bloc or stumbling bloc to global free trade in the presence of regionalism.[7] We isolate the effects of multilateralism by comparing the outcome of a world where multilateralism and regionalism both exist with a world where only regionalism exists.

In a comprehensive survey, Freund and Ornelas (2010, p. 156) document the "… scarcity of analyses on how multilateralism affects regionalism". Freund (2000) highlights how regionalism may follow from the success of multilateralism because an exogenous fall in multilateral tariffs can make an arbitrarily chosen bilateral FTA self-enforcing (when it is not so otherwise).[8] However, Freund abstracts from issues surrounding FTA proliferation. To focus on the FTA proliferation issue, we abstract from issues related to the self-enforcing nature of trade agreements and assume country pairs form FTAs whenever, anticipating any subsequent proliferation of FTAs, it is jointly optimal. Further, rather than take exogenous multilateral tariffs, we endogenize multilateral negotiations (and FTA formation). In doing so, we find multilateralism is never necessary for FTA formation and, indeed, the success of multilateralism is actually the reason it may prevent FTA expansion to global free trade.

Ornelas (2008) also investigates the link from multilateralism to regionalism, modeling multilateral negotiations before and after an arbitrary bilateral trade agreement. He shows world welfare rises upon FTA formation because of tariff complementarity, but an FTA does not emerge in equilibrium. Conversely, we find FTA formation emerges in equilibrium yet tariff complementarity may not emerge. We expand upon these differences in Section 5.

Our paper also links with other important papers in the broader trade agreements literature. In a three country setting, Bagwell and Staiger (2005b) analyze how rules, particularly non-discrimination and reciprocity, affect bilateral incentives to reduce tariffs after global negotiations. However, as the authors acknowledge, they abstract from the fact that these incentives really depend on whether the non-member to a bilateral agreement would form any subsequent agreements. We address this issue directly by modeling global negotiations among forward looking governments that correctly anticipate the extent of subsequent FTA formation. Indeed, as discussed above, globally negotiated tariff bindings not only affect the extent of FTA formation but the extent of FTA formation also affects the globally negotiated tariff bindings.[9] Our analysis also differs from Bagwell and Staiger (2005b) because our focus is isolating the role played by global negotiations in attaining global free trade by comparing the outcomes in the presence and absence of global negotiations.

Many papers in the literature emphasize a positive role for multilateral cooperation. In addition to Bagwell and Staiger (2005b), Maggi (1999) shows multilateralism can play a positive role by monitoring and punishing defectors. In contrast, our model shows how the presence of multilateral cooperation prior to bilateral cooperation can reduce world welfare.

Our paper also sheds light on the different empirical results of Estevadeordal et al. (2008) versus Limão (2006) and Karacaovali and Limão (2008). The former find empirical evidence for tariff complementarity among South American FTA members. However, the latter find no evidence that preferential tariff liberalization begets multilateral tariff liberalization for the US and the EU. Our theoretical results suggest the former (latter) should emerge among governments with relatively strong (weak) political economy motivations. Indeed, these predictions based on political economy motivations square well with the recent cross-country empirical estimates of political economy motivations by Gawande et al. (2012, 2015).

The binding overhang literature (i.e. globally negotiated tariff bindings exceeding applied tariffs) has two main explanations for its presence in an optimal trade agreement. First, Horn et al. (2010) argue costly contracting prevents formation of a state contingent global trade agreement. Second, many authors (see Bagwell and Staiger, 2005a , Amador and Bagwell, 2013 and Beshkar et al., 2015) argue that uncertainty over governments' future political economy motivations during global negotiations creates demand for flexibility over future applied tariffs.[10] Our explanation of binding overhang takes as given the practical observation that globally negotiated tariff bindings are not conditioned on a country's subsequent FTA formation behavior. Yet, the presence of multilateral tariff complementarity implies governments would like to condition tariff bindings in this way. Thus, the uncertainty in our model about which countries will subsequently form FTAs (a plausible situation during the 1994 Uruguay Round) creates a veil of ignorance and produces global tariff bindings whereby binding overhang can emerge because FTA members may still practice tariff complementarity. Section 5 discusses empirical differences relative to Beshkar et al. (2015).

[6] Prominent examples include Levy (1997), Krishna (1998) and Ornelas (2005a). More recent examples include Saggi and Yildiz (2010) and Lake (2017). See Freund and Ornelas (2010) for a recent extensive review.

[7] In doing so, our approach is closer to a strand of the literature beginning with Riezman (1999) that investigates the effect of FTA formation on the attainment of global free trade in a world where the only prevailing mechanism for trade liberalization is global tariff negotiations. Subsequent examples taking this perspective include Aghion et al. (2007), Saggi and Yildiz (2010) and Lake (2017).

[8] Similarly, Ethier (1998) argues regionalism is a benign consequence emerging from the success of multilateralism; it allows small countries that do not participate in early rounds of multilateral negotiations to gain by forming FTAs with large countries and attracting new foreign direct investment.

[9] When comparing our results to Bagwell and Staiger (2005b), one should keep in mind that our analysis implicitly embodies three rules: (i) complete bilateral tariff reductions, (ii) given symmetry, reciprocal and equal changes in member trade flows, and (iii) as FTA members maintain tariffs on the non-member, discriminatory bilateral tariff cuts.

[10] Private information over these motivations prevents a state contingent global trade agreement.

## 3. Model

### 3.1. Basic trade model

We use a competing exporters model, very similar to Bagwell and Staiger (1999b). There are three symmetric countries denoted by $i = a, b, c$ and three non-numeraire goods denoted by $Z = A, B, C$. Country $i$ has an endowment $e_i^Z = e$ of goods $Z \neq I$ and an endowment $e_i^Z = d < e$ of good $Z = I$. Below, we will see that country $i$ is a natural exporter of goods $Z \neq I$ and a natural importer of good $Z = I$. Thus, countries $j$ and $k$ are competing exporters in serving country $i$'s market. In turn, good $I$ can be viewed as country $i$'s "comparative disadvantage" good and goods $Z \neq I$ can be viewed as country $i$'s "comparative advantage" goods. Later, the hybrid parameter

$$\varphi \equiv \frac{e-d}{d}$$

appears frequently and represents the "strength of comparative advantage".

Given consumption $q^Z$ of each non-numeraire good $Z$ and $q^0$ of a numeraire good, consumer preferences are represented by $q^0 + \sum_{Z=A,B,C} u(q^Z)$ with the quasi-linearity implying the numeraire sector absorbs all general equilibrium effects. We assume country $i$'s demand for good $Z$ is $q^Z = q(p_i^Z) = \alpha - p_i^Z$ where $p_i^Z$ denotes the price of good $Z$ in country $i$. In turn, no arbitrage conditions link cross-country prices. Given non-prohibitive tariffs $t_{ij}$ and $t_{ik}$ applied by country $i$ on countries $j$ and $k$, $p_i^I = p_j^I + t_{ij} = p_k^I + t_{ik}$. Closed form solutions for domestic prices follow from international market clearing conditions. Letting $x_i^Z = e_i^Z - q(p_i^Z)$ denote country $i$'s *net* exports of good $Z$, market clearing for good $Z$ requires $\sum_i x_i^Z = 0$. Equilibrium domestic prices in country $i$ are then

$$p_i^I(t_{ij}, t_{ik}) = \alpha - \frac{1}{3}\left[(d+2e) - (t_{ij} + t_{ik})\right] \tag{1}$$

$$p_i^Z(t_{zi}, t_{zj}) = \alpha - \frac{1}{3}\left[(d+2e) - (t_{zj} - 2t_{zi})\right] \text{ for } Z \neq I. \tag{2}$$

Given these prices, country $i$'s net exports of good $Z \neq I$ to country $z \neq i$ are

$$x_{iz}^Z(t_{zi}, t_{zj}) = \frac{1}{3}\left[(e-d) + (t_{zj} - 2t_{zi})\right].$$

Thus, country $i$ is a natural exporter of goods $Z \neq I$ because $e > d$ implies $x_{iz}^Z(t_{zi}, t_{zj}) > 0$ when $t_{zi} = t_{zj} = 0$. Conversely, country $i$'s net imports of good $I$ from other countries are

$$-x_i^I(t_{ij}, t_{ik}) = \sum_{z=j,k} x_{zi}^I(t_{ij}, t_{ik}) = \tfrac{1}{3}\left[2(e-d) - (t_{ij} + t_{ik})\right].$$

Thus, country $i$ is a natural importer of good $I$ because $e > d$ implies $-x_i^I(t_{ij}, t_{ik}) > 0$ when $t_{ij} = t_{ik} = 0$. Moreover, $t_{jk} = t_{kj} = 0$ implies country $i$ has positive net exports of good $Z$ to country $z$ if and only if $t_{zi}$ falls below the prohibitive tariff

$$t_{PRO} \equiv \frac{1}{2}(e-d). \tag{3}$$

Thus, $t_{zi} < t_{PRO}$ preserves the competing exporters structure of the model.

It is well known that the effective partial equilibrium nature of the model implies country $i$'s national welfare can simply be represented as

$$W_i(\tau) = \sum_Z CS_i^Z(\tau) + \sum_Z PS_i^Z(\tau) + TR_i(\tau)$$

where $\tau \equiv (t_{ij}, t_{ik}, t_{ji}, t_{jk}, t_{ki}, t_{kj})$ is the global tariff vector, $CS_i^Z$ and $PS_i^Z$ denote country $i$'s consumer surplus and producer surplus associated with good $Z$ and $TR_i$ denotes country $i$'s tariff revenue. Appendix A contains algebraic expressions for the individual components of $W_i(\cdot)$. In addition to national welfare, the government's objective function in each country includes a political economy consideration based on the political influence emanating from the import competing sector. In particular, the payoff of country $i$'s government is given by

$$G_i(\tau) = \sum_Z CS_i^Z(\tau) + \sum_{Z \neq I} PS_i^Z(\tau) + (1+b)PS_i^I(\tau) + TR_i(\tau) \tag{4}$$

where $b > 0$ reflects the extent to which the government values protection of the import competing sector. To ensure optimal tariffs imposed by governments fall below the prohibitive tariff given by Eq. (3), we impose the following restriction hereafter:

$$b < \frac{1}{3}\varphi. \tag{5}$$

At this stage, it is useful to emphasize the role played by political economy motivations. As shown later by Eqs. (13)–(15), political economy motivations are the *only* reason governments negotiate non-zero tariffs during global negotiations. This should not be surprising given the literature recognizes that terms of trade externalities and political economy motivations are the two fundamental reasons why countries levy non-zero tariffs and that multilateral agreements neutralize terms of trade externalities (e.g. Bagwell and Staiger, 1999a). Thus, technically, political economy motivations allow us to model global tariff negotiations.

Nevertheless, one may question the economic relevance of political economy motivations given an important theme of the empirical Protection for Sale literature (e.g. Goldberg and Maggi, 1999 and Gawande and Bandyopadhyay, 2000) is that governments hold surprisingly weak political economy motivations. However, our main results are not inconsistent with this view as they rely on these motivations not being too strong. Nevertheless, we believe such motivations are empirically important determinants of tariffs. Indeed, recent contributions to the empirical Protection for Sale literature (e.g. Gawande et al., 2012, 2015) emphasize that governments have non-trivial political economy motivations upon recognizing (i) governments are influenced by both high tariff and low tariff interest groups and/or (ii) formally dealing with outliers in the data.

### 3.2. Global tariff negotiations and FTA negotiations

We adopt a simple, but flexible, protocol governing global tariff negotiations and FTA negotiations. We isolate the role that global tariff negotiations play by comparing the equilibrium outcomes of FTA negotiations that take place in the absence of global tariff negotiations versus after global tariff negotiations. Apart from the presence or absence of an initial round of global tariff negotiations (Stage 0), these two FTA formation games (Stages 1–3) are identical. Reflecting real world global tariff negotiations (e.g. Uruguay round), we model global negotiations over the upper bound on tariffs (i.e. tariff bindings) rather than actual tariffs (i.e. applied tariffs). Because countries are completely symmetric during global negotiations, we assume countries are treated symmetrically and model a common tariff binding.[11] Note, "binding overhang" can emerge because countries

[11] Section 6.1 extends our baseline analysis to include an asymmetric FTA negotiations protocol. This makes countries asymmetric at the global negotiations stage and, thus, we deal with asymmetric tariff bindings.

may set applied tariffs below the tariff binding after FTA negotiations conclude.

**Stage 0: Global negotiations.** Governments set the tariff binding cooperatively to maximize their joint expected payoff. To be clear, governments anticipate how the negotiated tariff bindings affect the equilibrium outcome of subsequent FTA negotiations.

**Stage 1: FTA negotiations.** Nature chooses whether or not FTA negotiations occur and, if so, the sequential order that country pairs negotiate FTAs. With probability $p \in (0, 1]$, FTA negotiations occur in Stages 1(a)–(c). As for the sequential order that country pairs negotiate FTAs, all six possible orderings are equally likely. When a country pair negotiates an FTA, each government of this "active pair" simultaneously announces whether or not to join an FTA with the other country in the active pair. An FTA forms if and only if both governments in the active pair choose to join the FTA. In the proofs, $a_i \in \{J, NJ\}$ denotes whether country $i$, as a member of an active pair, announces to join ($J$) or not join ($NJ$) an FTA with the other country in the active pair. With probability $1 - p$ there are no FTA negotiations, and thus no FTAs, and we move directly to the tariff setting stage (Stage 2).

**Stage 1(a).** Given the order previously chosen by nature, the three country pairs negotiate FTAs sequentially with the outcome of each pair's negotiation observed by all countries. However, once the first FTA forms, the game moves to Stage 1(b). If all three pairs fail to form an FTA, FTA negotiations end and the game moves directly to tariff setting (Stage 2).

**Stage 1(b).** Given the ordering chosen by nature, the two pairs who have not formed an FTA sequentially negotiate FTAs (even if they *chose* not to form an FTA in Stage 1(a)). However, once either pair forms an FTA, the game moves to Stage 1(c). If both pairs fail to form an FTA, the game moves directly to tariff setting (Stage 2).

**Stage 1(c).** The pair of countries yet to form an FTA has the opportunity to do so. Regardless of the outcome, the game moves to tariff setting (Stage 2).

Before describing tariff setting in Stage 2, note a desirable feature of our protocol: every pair of countries that chooses to not form an FTA in a given sub-stage gets a chance to reconsider in a later sub-stage if some other country pair forms an FTA. That is, FTA negotiations cease if and only if no pair of countries wants to form an additional FTA.[12,13]

**Stage 2: Tariff setting.** Governments of all countries choose their applied tariffs simultaneously subject to the zero tariff constraint between FTA members (GATT Article XXIV), the MFN principle (GATT Article I) and any prior globally negotiated tariff bindings.

**Stage 3: Production and consumption.** The applied tariffs set in Stage 2 determine production, trade, consumption and country payoffs $G_i(\tau)$.

Using backward induction, we solve for a pure strategy subgame perfect equilibrium of the FTA formation game. In doing so, we restrict attention to subgame perfect equilibria where FTA negotiations are efficient in the sense that when any pair of countries has an opportunity to form an FTA, they do so when mutually beneficial; this rules out equilibria where coordination failures prevent FTA formation.[14]

We will compare the equilibrium outcome of the FTA formation game when global tariff negotiations take place prior to the FTA formation game and the equilibrium outcome of the FTA formation game without global tariff negotiations. In particular, when global tariff negotiations precede the FTA formation game, the applied tariffs that countries set in Stage 2 are constrained by the globally negotiated tariff binding. However, in the absence of global tariff negotiations, the applied tariffs in Stage 2 are not bound by pre-existing tariff bindings since such bindings do not exist. Otherwise, the two FTA formation games are identical.

Importantly, our main results hold when FTA negotiations take place with certainty following global negotiations (i.e. $p = 1$). However, given FTA formation was relatively rare prior to the 1994 Uruguay Round of global negotiations, it is not clear whether governments perceived the subsequent flood of FTAs as likely or unlikely. Thus, the parameter $p$ captures the potential uncertainty regarding subsequent FTA formation in a simple way. In turn, we can perform comparative static exercises with $p$ and thereby investigate how government perception regarding the likelihood of future FTA negotiations affects the globally negotiated tariffs and the eventual extent of FTA formation.

Before examining optimal tariffs, we present a lemma underlying our analysis. The lemma deals with the incentive of countries who are the only country pair yet to form an FTA (i.e. Stage 1(c) of the FTA formation game). Hereafter, we denote an arbitrary network of FTAs by $g$ with the possible networks being: (i) no FTAs, $g = \emptyset$; (ii) a single FTA between countries $i$ and $j$, $g = g_{ij}$; (iii) two FTAs where country $i$ is the "hub" who is a member of both FTAs and the other countries $j$ and $k$ are "spokes", $g = g_i^H$; and (iv) global free trade, $g = g^{FT}$. $G_i(g)$ denotes government $i$'s payoff given the network $g$.

**Lemma 1.** *$G_i\left(g^{FT}\right) > G_i\left(g_j^H\right)$ so that spoke countries always form the FTA leading to global free trade. This is independent of whether global trade negotiations preceded FTA formation and any (non-zero) negotiated tariff binding.*

For spokes, the net benefit they obtain from FTA formation is weakly positive and proportional to the tariff they face in each others market. Given the hub has tariff free access to each spoke market, three reasons drive the attractiveness of spoke-spoke FTAs. First, the benefit of market access gained is high through eliminating the discrimination spokes face when exporting to each other. Second, the cost of domestic market access given up is low because the import competing sector's protection has already been diluted by the FTA with the hub. Third, given spokes already have an FTA with the hub, spoke-spoke FTAs are devoid of tariff complementarity and the associated intra-FTA negative externality.

### 3.3. Optimal tariffs

#### 3.3.1. Individually optimal tariffs

We now describe the individually optimal (i.e. non-cooperative) tariffs that countries set when unconstrained by tariff bindings.[15] They are important for solving the equilibrium structure of FTAs in the game without global tariff negotiations. However, they also play a role in the game with global tariff negotiations because, in general, the globally negotiated tariff binding may or may not exceed a country's individually optimal tariff and this determines whether the tariff binding actually constrains applied tariffs. Some tariff notation will only be used in the proofs with this notation explained at the beginning of Appendix B.

[12] The maximum number of FTA formation opportunities in Stage 1 is six. Stage 1(a) has a maximum of three, Stage 1(b) has a maximum of two and Stage 1(c) has only one.

[13] This feature makes the protocol more flexible than that in Aghion et al. (2007) where a single "leader" country can make sequential FTA proposals to two "follower" countries and the follower countries never have the opportunity to form their own FTA.

[14] We assume a country chooses not to join an FTA when indifferent between joining and not joining.

[15] These tariffs are all easily derived given the welfare expressions in Appendix A. In the special case of $b = d = 0$, they reduce to those found in Saggi and Yildiz (2010).

Given our government payoff expression $G_i(\cdot)$ and letting $x^I_{ii}$ denote output of good $I$ supplied by country $i$ to its domestic market, the first order condition (FOC) for $t_{ik}$ is:

$$\frac{\partial G_i(g)}{\partial t_{ik}} = \left[\left(1 - \frac{\partial p^I_i}{\partial t_{ik}}\right)x^I_{ki} - \frac{\partial p^I_i}{\partial t_{ik}}x^I_{ji}\right] + \left[t_{ik}\frac{\partial x^I_{ki}}{\partial t_{ik}} + t_{ij}\frac{\partial x^I_{ji}}{\partial t_{ik}}\right] + \left[bx^I_{ii}\frac{\partial p^I_i}{\partial t_{ik}}\right]. \tag{6}$$

Following Ornelas (2005b), we refer to the three terms in square brackets as, respectively, the (i) terms of trade effect, (ii) tariff revenue effect, and (iii) distributive effect.[16] In general, country $i$ depresses the world price and increases the tariff inclusive domestic price of its imported good $I$ by imposing tariffs. However, when only raising $t_{ik}$, country $i$'s terms of trade improve vis a vis country $k$ (i.e. $1 - \frac{\partial p^I_i}{\partial t_{ik}} > 0$) but deteriorate vis a vis country $j$ (i.e. $-\frac{\partial p^I_i}{\partial t_{ik}} < 0$) because country $j$ now receives the higher tariff inclusive domestic price when exporting to country $i$ and faces an unchanged tariff $t_{ij}$. The tariff $t_{ik}$ also affects tariff revenue by reducing imports and shifting the composition of imports away from country $k$ and towards country $j$ $\left(\frac{\partial x^I_{ki}}{\partial t_{ik}} > \frac{\partial x^I_{ji}}{\partial t_{ik}} > 0\right)$.[17] Finally, the distributive effect captures the redistribution of domestic surplus from consumers to producers which is valuable given the government's political motivations.

Absent FTAs, solving the FOCs for the tariffs imposed by country $i$'s government on countries $j$ and $k$, i.e. $t_{ij}(\emptyset)$ and $t_{ik}(\emptyset)$, yields:

$$t_{ij}(\emptyset) = t_{ik}(\emptyset) \equiv t_{Nash} = \frac{1}{4}(e-d) + \frac{3}{4}bd. \tag{7}$$

Country $i$ chooses non-discriminatory tariffs because of symmetry with these tariffs consisting of two terms. The first term reflects the terms of trade and tariff revenue motives in the absence of political economy motivations. In particular, larger domestic import competing sectors (i.e. higher $d$) reduce world export volumes and, in turn, mitigate the terms of trade motive. The second term reflects the influence of government political economy motivations that emerge directly via the distributive effect and also indirectly via the impact of politically charged tariffs on the terms of trade and tariff revenue effects. Naturally, the political economy influence strengthens with the extra weight placed on the import competing sector's producer surplus, $b$, and the size of the domestic import competing sector, $d$.[18,19] Fig. 1 illustrates various tariffs discussed in this section.

We now describe how FTA formation affects optimal tariffs. First, FTA formation between countries $i$ and $j$ (insiders) leaves the optimal tariffs of country $k$ (outsider) unchanged:

$$t_{ki}(g_{ij}) \equiv t^*_{OUT} = \frac{1}{4}(e-d) + \frac{3}{4}bd = t_{Nash}. \tag{8}$$

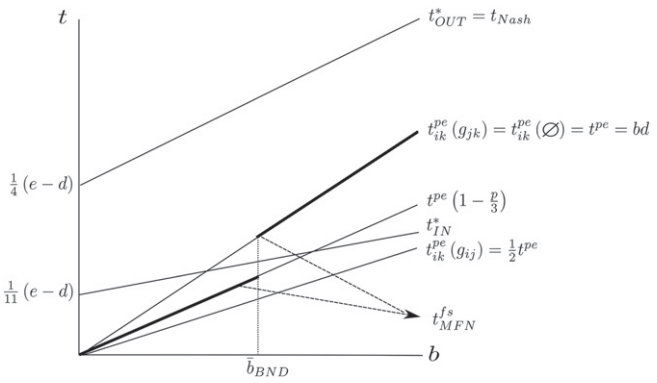


**Fig. 1.** Individually optimal and jointly optimal tariffs

This follows from the separability of goods markets which implies $k$'s incentive to manipulate the price of its imported good is independent of the tariffs on other goods and, indeed, an FTA between $i$ and $j$ affects the tariffs on these other goods. Moreover, in our model, the outsider government's political economy motivations depend exclusively on the market of its imported good and thus do not depend on tariffs for other goods.

Second, FTA insiders choose to lower their tariff on the non-member outsider, a phenomenon known as tariff complementarity. Hereafter, we refer to it as "individual" tariff complementarity. An insider's optimal tariff (say country $i$) on the outsider country $k$ is

$$t_{ik}(g_{ij}) \equiv \frac{1}{11}(e-d) + \frac{3}{11}bd \equiv t^*_{IN}. \tag{9}$$

Individual tariff complementarity follows from $t^*_{IN} < t_{Nash} = t^*_{OUT}$. Intuitively, the FTA between countries $i$ and $j$ weakens the terms of trade and tariff revenue motivations for country $i$'s external tariff on country $k$. The underlying cause is that the FTA shifts the composition of $i$'s imports towards country $j$. When raising $t_{ik}$, the importance of country $i$'s terms of trade deterioration vis a vis country $j$ rises while the importance of its terms of trade improvement vis a vis country $k$ falls. Moreover, country $i$'s ability to raise tariff revenue from the non-member $k$ falls. Thus, weaker terms of trade and tariff revenue motivations of country $i$ explain the individual tariff complementarity effect.[20]

Finally, as above, formation of a second FTA between, say, countries $i$ and $k$ leaves the tariff of the non-member, country $j$, unaffected: $t_{jk}(g^H_i) = t_{jk}(g_{ij})$. However, as above, the outsider country $k$ lowers its tariff on the non-member country $j$ so that:[21]

$$t_{kj}\left(g^H_i\right) = \frac{1}{11}(e-d) + \frac{3}{11}bd = t^*_{IN}. \tag{10}$$

### 3.3.2. Optimal globally negotiated tariff bindings

We now describe the jointly optimal tariff binding that governments negotiate prior to FTA formation. Due to symmetry, we naturally assume that governments maximize their joint payoff. Moreover, given the independence of markets, we merely focus on the jointly optimal tariff in the market of good $I$ which is imported by country $i$. For the sake of exposition, we initially assume governments negotiate future applied tariffs imposed by countries and can

[16] Ornelas' general setup also includes a fourth term $(1+b)p^I_i\frac{\partial x^I_{ii}}{\partial t_{ik}}$ which he labels the strategic effect. However, $\frac{\partial x^I_{ii}}{\partial t_{ik}} = 0$ in our model because of the endowment structure.

[17] In a completely symmetric setting, the terms of trade and distributive effects are positive while the tariff revenue effect is negative. This follows upon letting $t_{ik} = t_{ij}$ and $x^I_{ki} = x^I_{ji}$.

[18] Note that our assumption in Eq. (5) on the range of the parameter $b$ implies that the Nash tariffs are below the prohibitive level $t_{PRO}$ given in Eq. (3).

[19] Although we assume symmetric political preferences, the effect of $b$ on an importing country's tariff (whether individually optimal or jointly optimal) always has the interpretation of the country's own political preference. This follows from the separability of goods markets: country $j$'s tariff on its imported good (which depends on country $j$'s political preference) does not affect the market for the good imported by country $i$ and hence does not affect country $i$'s tariff.

[20] Note that the distributive effect, $bx^I_{ii}\frac{\partial p^I_i}{\partial t_{ik}}$, is independent of $t_{ij}$ in our model and so the *only* reason tariff complementarity emerges is because of the effects of the FTA between $i$ and $j$ on the terms of trade and tariff revenue motives.

[21] Of course, since the hub country has FTAs with both of the other countries then it practices free trade.

condition these applied tariffs on whether a country has formed FTAs or not. Naturally, given these assumptions contradict real world tariff setting, we relax these assumptions when determining the optimal tariff bindings.

Letting $G^I(g;(t_{ij},t_{ik})) = \sum_{z=a,b,c} G^I_z(g;(t_{ij},t_{ik}))$ denote the joint government payoff in market $I$ when the network of FTAs is $g$, governments maximize their joint payoff by solving:

$$\max_{t_{ij},t_{ik}} G^I(g;(t_{ij},t_{ik})). \tag{11}$$

In our model, the FOC for $t_{ik}$ is given by:

$$bx^I_{ii}\frac{\partial p^I_i}{\partial t_{ik}} + \left[t_{ik}\frac{\partial x^I_{ki}}{\partial t_{ik}} + t_{ij}\frac{\partial x^I_{ji}}{\partial t_{ik}}\right] = 0. \tag{12}$$

When comparing this FOC for governments' jointly optimal $t_{ik}$ and the FOC for the individually optimal $t_{ik}$ in Eq. (6), three observations stand out. First, as is well known, the jointly optimal tariff bears no imprint of the terms of trade effects that enter country $i$'s individually optimal tariff. Second, the two terms in Eq. (12) shaping the jointly optimal $t_{ik}$ are the distributive and tariff revenue effects present in country $i$'s individually optimal $t_{ik}$. These two observations imply the only difference between the incentives underlying the jointly optimal and individually optimal $t_{ik}$ is that terms of trade motivations do not impact the jointly optimal $t_{ik}$. In turn, the third observation is that the individually optimal $t_{ik}$ is less sensitive to a rising $b$ than the jointly optimal $t_{ik}$. Specifically, the terms of trade motive weakens as $b$ rises because tariff levels rise with stronger political economy motivations which depresses world export volumes and, hence, the terms of trade motive. Thus, the individually optimal $t_{ik}$ is less sensitive to a rising $b$ than the jointly optimal $t_{ik}$.

Absent FTAs, solving the FOC Eq. (12) for $t_{ik}$ and an analogous FOC for $t_{ij}$ reveals the jointly optimal tariffs. We refer to these as "politically efficient" tariffs and they are given by the non-discriminatory tariffs:

$$t^{pe}_{ij}(\emptyset) = t^{pe}_{ik}(\emptyset) = bd \equiv t^{pe}. \tag{13}$$

Given the separability of markets, these politically efficient tariffs in the absence of FTAs are also the politically efficient tariffs for an outsider:

$$t^{pe}_{ij}(g_{jk}) = t^{pe}_{ik}(g_{jk}) = t^{pe}. \tag{14}$$

However, FTA formation affects the politically efficient tariff for insiders. When countries $i$ and $j$ form an FTA, solving the FOC Eq.(12) after imposing $t_{ij} = 0$ reveals

$$t^{pe}_{ik}(g_{ij}) = \frac{1}{2}bd = \frac{1}{2}t^{pe}. \tag{15}$$

The lower politically efficient tariff for an insider upon FTA formation, i.e. $t^{pe}_{ik}(g_{ij}) < t^{pe}_{ik}(\emptyset)$, indicates the presence of "multilateral" tariff complementarity, identified by Ornelas (2008). Given our discussion surrounding the FOC Eq. (12), multilateral tariff complementarity emerges because the tariff revenue effect still enters the jointly optimal tariff for an insider.

Our analysis above assumed that governments negotiate applied tariffs and can condition future applied tariffs on the structure of FTAs. In practice, governments negotiate tariff bindings rather than applied tariffs and do not condition future tariff bindings of a country on its future formation of FTAs. We now incorporate these two realities. In particular, governments negotiate the global tariff binding anticipating that FTA formation could subsequently occur but without knowing who would form such FTAs. In this case, Lemma 2 characterizes the optimal tariff binding when countries anticipate a single FTA will subsequently emerge and Fig. 1 helps illustrate graphically. Note, external tariffs refer to applied tariffs apart from the zero applied tariffs between FTA members.

**Lemma 2.** *Suppose that governments anticipate a single FTA will emerge if FTA negotiations take place. Then, there exists a critical value of b, denoted* $\bar{b}_{BND} > 0$*, such that global negotiations lead to the following optimal tariff binding* $t^{fs}_{MFN}$*:*

$$t^{fs}_{MFN} \equiv \begin{cases} t^{pe}\left(1-\frac{p}{3}\right) & \text{if } b < \bar{b}_{BND} \\ t^{pe} & \text{if } b \geq \bar{b}_{BND} \end{cases}.$$

*External tariffs are bound by* $t^{fs}_{MFN}$ *except when* $b \geq \bar{b}_{BND}$ *so that insiders set* $t^*_{IN} < t^{fs}_{MFN}$.

When governments anticipate subsequent formation of a single FTA conditional on FTA negotiations taking place, the jointly optimal tariff imposed by country $i$ reflects that country $i$ could be an insider or an outsider (with respective probabilities $\frac{2}{3}$ and $\frac{1}{3}$) and that FTA negotiations may or may not take place (with respective probabilities $p$ and $1-p$). Recognizing these uncertainties, the optimal binding that binds the insiders and the outsider is the farsighted MFN tariff $t^{fs}_{MFN} = t^{pe}\left(1-\frac{p}{3}\right)$ which solves

$$\begin{aligned}\arg\max_t \quad & p\frac{1}{3}\left[G^I(g_{ij};(0,t_{ik}=t)) + G^I(g_{ik};(t_{ij}=t,0))\right. \\ & \left. + G^I(g_{jk};(t_{ik}=t,t_{jk}=t))\right] + (1-p)\,G^I(\emptyset;(t_{ij}=t,t_{ik}=t)). \end{aligned} \tag{16}$$

Two fundamental motives drive the tariff binding concessions embodied in the farsighted MFN tariff. Note that $t^{fs}_{MFN} = t^{pe}\left(1-\frac{p}{3}\right)$ is an "average" politically efficient tariff that averages over (i) the politically efficient tariffs for the insider and the outsider and (ii) whether FTA negotiations take place or not:

$$\begin{aligned} t^{pe}\left(1-\frac{p}{3}\right) &= p\left[\frac{2}{3}t^{pe}_{ik}(g_{ij}) + \frac{1}{3}t^{pe}_{ik}(g_{jk})\right] + (1-p)\,t^{pe}_{ik}(\emptyset) \\ &= \frac{2p}{3}t^{pe}_{ik}(g_{ij}) + \left(1-\frac{2p}{3}\right)t^{pe}_{ik}(\emptyset) \end{aligned}$$

where the last line follows from $t^{pe}_{ik}(g_{jk}) = t^{pe}_{ik}(\emptyset)$. Thus, one can decompose $t^{fs}_{MFN}$ into a first motive explaining why $t^{pe}_{ik}(g_{ij})$ differs from $t^{pe}_{ik}(\emptyset)$ and a second motive explaining $t^{pe}_{ik}(\emptyset)$. The former explanation is multilateral tariff complementarity. The latter explanation is that the politically efficient tariff $t^{pe}_{ik}(\emptyset)$ removes unilateral terms of trade imprints from individually optimal tariffs. Thus, tariff binding concessions reflect the global efficiency implications of multilateral tariff complementarity and unilateral terms of trade incentives.[22]

While $t^{fs}_{MFN} = t^{pe}\left(1-\frac{p}{3}\right)$ is the optimal binding conditional on binding the insiders and the outsider, governments could set a tariff binding that only binds the outsider upon FTA formation.[23] In this

[22] One may have suspected that the "average" politically efficient tariff reflects an insurance motive whereby individual governments want to smooth their payoff across the uncertainty about being an insider or an outsider. This is incorrect and starkly illustrated by Section 6.1.2 where we treat the identity of the insiders as known with certainty yet without any affect on the global tariff binding described here.

[23] Since tariff complementarity implies $t_{Nash} = t^*_{OUT} > t^*_{IN}$, it is not possible to set a tariff binding that only binds insiders. Moreover, setting a tariff binding that does not bind any applied tariffs is not optimal.

case, the optimal tariff binding for the outsider is $t_{ik}^{pe}(g_{jk}) = t^{pe}$ while insiders set their individually optimal tariff $t_{IN}^{*}$. The critical value $\bar{b}_{BND}$ determines whether governments find it optimal to bind the insiders and the outsider or only bind the outsider.[24]

Fig. 1 shows how $\bar{b}_{BND}$ balances the tension between the cost and benefit of binding the insiders and the outsider versus only binding the outsider. Binding the insiders and the outsider via a tariff binding $t^{pe}\left(1-\frac{p}{3}\right)$ is costly because the tariff imposed by the outsider falls below the politically efficient tariff for an outsider of $t_{ik}^{pe}(g_{jk}) = t^{pe}$. But, the benefit is that the tariff imposed by the insider falls from the individually optimal level $t_{IN}^{*}$ towards the politically efficient tariff for an insider of $t_{ik}^{pe}(g_{ij}) = \frac{1}{2}t^{pe}$. Crucially, as discussed above and illustrated in Fig. 1, individually optimal tariffs are less sensitive than politically efficient tariffs to a rising $b$ (because the terms of trade motive weakens as $b$ rises). When $b$ is low, the benefit of binding the insiders and the outsider is high while the cost is proportional to $b$ and, hence, small. But, as $b$ rises, the benefit of binding the insiders and the outsider falls (i.e. $t_{IN}^{*} - t^{pe}\left(1-\frac{p}{3}\right)$ shrinks) while the cost, which is proportional to $b$, rises. The critical value $\bar{b}_{BND}$ exactly balances the benefit and cost with governments choosing to bind the insiders and the outsider when $b < \bar{b}_{BND}$ but only bind the outsider when $b > \bar{b}_{BND}$.[25]

Before moving on, we note an important result of our model: the shadow of future FTA formation feeds into the initial globally negotiated tariff bindings as seen in Lemma 2.

## 4. Global tariff negotiations and global free trade

To assess the role played by global tariff negotiations in the attainment of global free trade, we first investigate the extent of FTA formation following global negotiations. While Section 5 characterizes *how many* FTAs form, our main priority now is *whether* FTA expansion leads to global free trade when global negotiations precede FTA formation.

Two results from the previous section provide the starting point. First, Lemma 1 says a hub-spoke network cannot emerge in equilibrium. Thus, FTA formation either stops at a single FTA or expands to global free trade. Second, Lemma 2 says implementing the farsighted MFN tariff $t_{MFN}^{fs}$ as the globally negotiated tariff binding maximizes the joint expected government payoff when, conditional on FTA negotiations taking place, a single FTA emerges in equilibrium. Thus, if governments anticipate a single FTA will emerge in equilibrium then they will implement the farsighted MFN tariff as the global tariff binding. The key question now is the following: what is the equilibrium outcome when governments implement the farsighted MFN tariff as the global tariff binding?

Lemma 3 states the answer.

**Lemma 3.** *Suppose governments set the farsighted MFN tariff $t_{MFN}^{fs}$ as the global tariff binding. (i) At most a single FTA forms in equilibrium. (ii) If $b < \bar{b}_{BND}$ then a single FTA forms in equilibrium when FTA negotiations take place. (iii) Governments' joint expected payoff at the global negotiations stage exceeds that under global free trade.*

While Lemma 3 says a single FTA is not necessarily the only equilibrium outcome when governments implement the farsighted MFN tariff as the global tariff binding, it says the only other possible outcome is no FTAs. Moreover, regardless of the equilibrium outcome, governments have a higher joint expected payoff than under global free trade.

Who resists expansion of a single FTA to global free trade after negotiating the farsighted MFN tariff as the global tariff binding? Naturally, foreseeing subsequent FTA formation eventually yields global free trade, an insider only engages in formation of a second FTA with the outsider if its eventual payoff under global free trade exceeds that as an insider. The main advantage an insider receives from global free trade is elimination of the tariff barrier faced when exporting to the outsider. However, this incentive is relatively weak given the global tariff binding $t_{MFN}^{fs}$ significantly restrains the outsider's applied tariff. Moreover, the insider's own political economy motivations further reduce the incentive to engage in subsequent FTA formation. As a result, the insider chooses not to form a second FTA and therefore blocks further FTA expansion. Thus, at most a single FTA emerges.

Indeed, a single FTA emerges in equilibrium when $b < \bar{b}_{BND}$ and governments set the farsighted MFN tariff $t_{MFN}^{fs}$ as the global tariff binding. Anticipating that a single FTA will not expand any further, the benefit a potential insider receives from *not* becoming an insider lies in the political benefit of maintaining protection for the import competing sector via the tariff imposed on the other potential insider. However, this political benefit is small when $b < \bar{b}_{BND}$ because the politically efficient tariff $t^{pe}\left(1-\frac{p}{3}\right)$ is already placing considerable restraint on the applied tariff of each potential insider. Thus, upon setting $t_{MFN}^{fs}$ as the global tariff binding, a single FTA emerges when $b < \bar{b}_{BND}$.

Regardless of whether a single FTA or no FTAs emerge, the joint expected government payoff at the global negotiations stage exceeds that under global free trade. This follows by construction when a single FTA emerges because the farsighted MFN tariff maximizes the joint expected government payoff conditional on a single FTA subsequently emerging. In particular, the joint expected government payoff exceeds that under global free trade as governments have the option of setting a zero tariff binding. Moreover, Lemma 3 says no FTAs can emerge only if $b > \bar{b}_{BND}$. But, in this case, the farsighted MFN tariff is the politically efficient tariff $t_{MFN}^{fs} = t_{ik}^{pe}(g_{jk}) = t_{ik}^{pe}(\emptyset) = t^{pe}$ and, by definition, the maximum joint payoff that governments can ever attain is when no FTAs form and global applied tariffs are given by $t^{pe}$. This discussion implies global free trade never emerges: governments have the option of setting the farsighted MFN tariff knowing such a tariff binding does not lead to global free trade and always delivers a higher joint expected payoff than global free trade. We state this important result in the following proposition.

**Proposition 1.** *Global free trade never emerges when global tariff negotiations take place prior to FTA negotiations.*

While global free trade never emerges in the presence of global tariff negotiations, establishing the role played by global tariff negotiations in the attainment of global free trade depends on whether global free trade would be attained in the absence of such negotiations. To establish the equilibrium in the absence of global tariff negotiations, we now consider the FTA formation game in the absence of global negotiations. Here, FTA members eliminate tariffs on each other but governments are not constrained by any pre-existing tariff bindings.

Unless political economy considerations are very strong, at least one FTA must form. In a world without FTAs, all applied tariffs would equal the non-cooperative Nash tariff $t_{Nash}$. As such, FTA formation would bring significant welfare gains to each member government that outweighs the political cost. Further, Lemma 1 says a hub-spoke network cannot emerge in equilibrium because spoke countries benefit by deviating and forming their own FTA that takes the world to global free trade. Thus, the equilibrium outcome in the absence of

[24] In the proof of Lemma 2 we establish that the farsighted MFN tariff actually binds all external tariffs when $b < \bar{b}_{BND}$ but only binds the outsider's external tariffs when $b \geq \bar{b}_{BND}$. See Eq. (21) in the proof of Lemma 2 for the algebraic expression of $\bar{b}_{BND}$.

[25] Note, governments are indifferent between setting $t^{pe}$ or $t^{pe}\left(1-\frac{p}{3}\right)$ as the tariff binding when $b = \bar{b}_{BND}$. We assume they set $t^{pe}$ when $b = \bar{b}_{BND}$.

global tariff negotiations must be either a single FTA or global free trade.

This brings us to the important issue of why the *absence* of global tariff negotiations can lead to global free trade as the equilibrium outcome rather than a fragmented world with only a single FTA. Both insiders and the outsider recognize formation of a second FTA eventually leads to global free trade. However, the relative attractiveness of global free trade differs for the insiders and the outsider. For all countries, global tariff elimination brings additional market access for exporters and reduced protection for the domestic import competing sector with the latter becoming more costly as political economy motivations strengthen. But the outsider reaps an additional gain because it no longer faces discrimination in the FTA member markets. Thus, if the tariff imposed by insiders on the outsider and that imposed by the outsider on the insiders are equal, then this "discrimination effect" implies that the outsider has a weaker incentive than the insider to block global free trade.

However, as discussed in Section 3.3, individual tariff complementarity lowers an insider's optimal tariff $t^*_{IN}$ imposed on the outsider below the outsider's optimal tariff $t^*_{OUT}$ imposed on the insider. Thus, the insider's import competing sector now loses less and the outsider's exporting sector now gains less upon expansion to global free trade. Indeed, these effects of tariff complementarity outweigh the discrimination effect so that the outsider has a stronger incentive to block global free trade. Put slightly differently, the absence of tariff concessions given by the outsider motivate each insider's desire to engage in subsequent FTA formation with the outsider even though it eventually yields global free trade. When interpreting our main result, this observation will be very important.

While the outsider has a stronger incentive to block global free trade, whether it does so depends on its political economy motivations. An outsider refuses participation in subsequent FTA formation, thereby blocking global free trade, when $G_i(g_{jk}) \geq G_i(g^{FT})$. Not surprisingly, given the optimal tariffs of insiders and outsiders discussed in Section 3.3, an outsider blocks global free trade only if political economy motivations exceed a threshold

$$b \geq \bar{b}_{OUT} \equiv \frac{13}{137}\varphi. \tag{17}$$

If $b < \bar{b}_{OUT}$, an outsider does not block global free trade and hence global free trade emerges in the absence of global tariff negotiations. In this case, FTA formation represents the only, albeit blunt, mechanism whereby insiders can extract tariff concessions from the outsider.[26]

Proposition 2 now presents our main result.

**Proposition 2.** *Global tariff negotiations prevent global free trade when $b < \bar{b}_{OUT}$ (where $\bar{b}_{OUT}$ is defined in* Eq. (17)*).*

Global tariff negotiations prevent global free trade because global free trade never emerges in the presence of global tariff negotiations (Proposition 1) yet emerges in the absence of global tariff negotiations when $b < \bar{b}_{OUT}$. In other words, global tariff negotiations are actually the cause of a world stuck short of global free trade when political economy motivations are "not too large". Notice that, given our parameter space is restricted to $b < \frac{1}{3}\varphi$, the striking result of Proposition 2 holds for nearly one-third of the parameter space. Moreover, given the parameter $\varphi$ can be arbitrarily large as $d$ approaches 0, the result in Proposition 2 may hold even when political economy motivations are very strong.

Gaining a better understanding of how global tariff negotiations prevent global free trade requires understanding how the presence of global negotiations change the incentives of the outsider or the insiders such that one of them now refuses participation in FTA expansion that would ultimately yield global free trade. As noted above, the insider opted against blocking global free trade in the absence of global tariff negotiations because it had not extracted any tariff concessions from the outsider. But, the presence of global tariff negotiations leads to a relatively low tariff binding and, as such, extracts significant applied tariff concessions from the eventual outsider. Indeed, these tariff concessions received by the eventual insider (through tariff bindings set by forward looking governments during global negotiations) are large enough that an insider now refuses participation in FTA expansion and, thus, blocks expansion to global free trade. Therefore, the role of tariff concessions given by the eventual outsider in global tariff negotiations drive the result that global tariff negotiations can prevent global free trade. More broadly, the success of global tariff negotiations in lowering tariff bindings and applied tariffs across all participating countries underlies why global tariff negotiations prevent global free trade.

## 5. A fragmented world of gated globalization

Section 4 established that global tariff negotiations prevent global free trade primarily because the negotiated tariff concessions eliminate the FTA expansion incentives necessary for global free trade to emerge via FTA formation. Although Lemmas 1–3 established that a single FTA or no FTAs must emerge in equilibrium following global tariff negotiations, we did not characterize the conditions governing whether global negotiations lead to a single FTA and a fragmented world of globalization or whether they yield a world of no FTAs. In particular, while Lemma 3 established a threshold level of political economy motivations that ensures no FTAs emerge upon setting $t^{fs}_{MFN}$ as the global tariff binding, is it possible that governments can and/or want to prevent FTA formation by not setting $t^{fs}_{MFN}$ as the global tariff binding? And, if so, what are the equilibrium global tariff bindings?

To begin, what tariff bindings make FTA formation unattractive for insiders relative to the absence of any FTAs (i.e. $G_i(g_{ij};\cdot) < G_i(\emptyset;\cdot)$) and, hence, prevent FTA formation? The answer depends on a trade-off between the welfare gains of an FTA and a government's desire to protect its import competing sector. In particular, governments must have sufficiently strong political economy motivations if they forego FTA formation opportunities.

A government's overall political economy motivations depend on the wedge between its payoff and national welfare which, as seen in Eq. (4), is $b \cdot PS^I_i$. Thus, a necessary condition for no FTA formation is that $b$ exceeds a threshold; specifically, $b \geq \frac{1}{8}\varphi$. For $b < \frac{1}{8}\varphi$, FTA formation cannot be deterred regardless of the global tariff binding. However, $b \geq \frac{1}{8}\varphi$ is not a sufficient condition. Insiders opt against becoming insiders only if the import competing sector is strong enough given that a government's overall political economy motives depend on the size of its producer surplus. Because higher tariffs strengthen the import competing sector, the tariff binding must be large enough. Thus, governments opt against FTA formation only if the tariff binding exceeds a threshold $\underline{t}(b)$ (see Eq. (24) in the Appendix for the algebraic expression) in addition to $b \geq \frac{1}{8}\varphi$. Lemma 4 summarizes this discussion.

**Lemma 4.** *For $b < \frac{1}{8}\varphi$, there are no global tariff bindings that prevent all FTA formation. For $b \geq \frac{1}{8}\varphi$, there exits a threshold $\underline{t}(b)$ such that a global tariff binding $t$ prevents all FTA formation only if $t \geq \underline{t}(b)$.*

[26] The effect of the symmetric $b$ on a country's incentive for FTA formation aggregates the separate effects stemming from each member's political preference. Note, a country's individually optimal tariff rises with $b$ and the value of market access gained or given increases with the tariff level. Thus, a country would prefer FTA formation with a higher $b$ parter and a country's benefit of FTA formation falls with its own $b$. The inequality in Eq. (17), and similar inequalities in the proof of Proposition 2, indicate the latter effect dominates. See Stoyanov and Yildiz (2015) for an analysis of FTA formation under asymmetric political preferences.

Given Lemma 4 establishes FTA formation takes place when $b < \frac{1}{8}\varphi$ regardless of the global tariff binding, we suppose hereafter that $b \geq \frac{1}{8}\varphi$. Nevertheless, under what conditions would governments jointly prefer deviating from the tariff binding $t^{fs}_{MFN}$ to some tariff binding above $\underline{t}(b)$ in order to prevent all FTAs?

If governments could pre-commit to abstain from FTA formation at the global negotiations stage, this would be jointly optimal. In doing so, they would set a tariff binding equal to the politically efficient tariff $t^{pe}_{ij}(\emptyset) = t^{pe}$ which would bind the applied tariffs of all countries. However, in reality and in our framework, governments cannot credibly make such commitments. Nevertheless, governments may be prepared to sacrifice some political efficiency in order to prevent FTA formation. Naturally, doing so becomes less attractive as governments move further away from the politically efficient tariff $t^{pe}_{ij}(\emptyset) = t^{pe}$. Thus, if governments can prevent FTAs through a tariff binding sufficiently close to the politically efficient tariff $t^{pe}_{ij}(\emptyset) = t^{pe}$ then doing so is jointly optimal; otherwise, they are better off staying with the tariff binding $t^{fs}_{MFN}$ and the single FTA outcome.

Specifically, governments jointly opt against preventing FTA formation if the minimum required tariff binding for prevention, given by $\underline{t}(b)$, exceeds $t^{pe} + x(b)$ (the algebraic expression for $x(b) > 0$ is given by Eq. (26) in the Appendix). Conversely, governments prevent FTA formation by setting a tariff binding equal to $\max\{\underline{t}(b), t^{pe}\}$ if $\underline{t}(b) < t^{pe} + x(b)$ because the associated sacrifice in political efficiency is small enough. Indeed, we can solve for a threshold value of political economy motivations $\bar{b}_{\emptyset}$ such that governments are indifferent between preventing and not preventing FTA formation:

$$t^{pe} + x(b) = \underline{t}(b) \text{ if and only if } b = \bar{b}_{\emptyset}. \tag{18}$$

The equilibrium characterization now follows easily in Proposition 3.

**Proposition 3.** *Global tariff negotiations lead to (i) a fragmented world with a single FTA when FTA negotiations take place and $b < \bar{b}_{\emptyset}$ but (ii) a world without FTAs when $b \geq \bar{b}_{\emptyset}$. Moreover, global negotiations produce a tariff binding $t^{fs}_{MFN}$ where*

$$t^{fs}_{MFN} = \begin{cases} t^{pe}\left(1 - \frac{p}{3}\right) & \text{if } b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\} \\ t^{pe} & \text{if } b \in \left[\bar{b}_{BND}, \bar{b}_{\emptyset}\right) \\ \max\{\underline{t}(b), t^{pe}\} & \text{if } b \geq \bar{b}_{\emptyset} \end{cases}.$$

*When $b < \bar{b}_{\emptyset}$, external tariffs are bound by $t^{fs}_{MFN}$ except for insiders when $b \in \left[\bar{b}_{BND}, \bar{b}_{\emptyset}\right)$ in which case they set $t^{*}_{IN} < t^{fs}_{MFN}$.*

Fig. 2 illustrates Proposition 3. When FTA negotiations take place, a single FTA emerges if and only if political economy motivations fall below $\bar{b}_{\emptyset}$. When $b < \bar{b}_{\emptyset}$, the sacrifice of political efficiency needed to prevent FTA formation is too large (i.e. $\underline{t}(b) > t^{pe} + x(b)$). In turn, governments set the tariff binding equal to $t^{fs}_{MFN}$ and a single FTA emerges (if FTA negotiations occur). Further, as discussed above, this tariff binding binds all external tariffs except when $b \geq \bar{b}_{BND}$. In this case, $t^{fs}_{MFN} = t^{pe}$ and insiders lower their applied tariff on the outsider from $t^{pe}$ to $t^{*}_{IN} < t^{pe}$ upon FTA formation. However, governments prevent FTA formation once $b \geq \bar{b}_{\emptyset}$ by setting the tariff binding $\underline{t}(b)$ or, once $b > b^{*}$, $t^{pe}$. When setting $\underline{t}(b)$, the sacrifice in political efficiency is small enough that governments set the tariff bindings away from the politically efficient tariff $t^{pe}_{ij}(\emptyset) = t^{pe}$ to prevent FTA formation.

Our gated globalization result in Proposition 3 differs from Ornelas (2008). Assuming governments (i) know the identity of insiders and the outsider and (ii) Nash bargain over multilateral tariffs, Ornelas finds that FTA formation cannot emerge in equilibrium

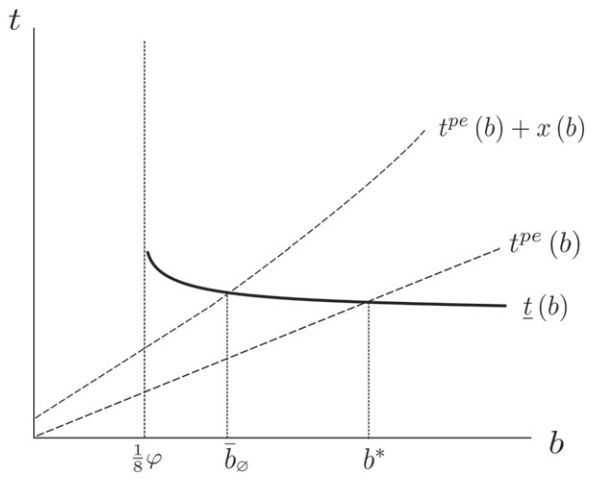


**Fig. 2.** When does a single FTA arise in equilibrium?

because individually tariff complementarity substantially improves the outside option of the FTA outsider. On the surface, numerous explanations could reconcile these results. Unlike Ornelas, our baseline analysis assumes that, during global negotiations, (i) the identity of insiders and the outsider is unknown and (ii) an individual country cannot use the subsequent FTA formation outcome to extract greater concessions. But, upon relaxing these assumptions in Section 6.1, a single FTA still emerges. The key explanation is that, unlike the Nash bargaining assumption of Ornelas, our countries do not equally split the joint surplus created via multilateral cooperation. For example, the farsighted MFN tariff maximizes governments joint expected payoff but, relative to the global free trade payoff, raises an eventual insider's payoff and lowers the eventual outsider's payoff. If the outsider's identity were known and it could withdraw from global negotiations, as in Section 6.1.2, the insiders would concede a tariff binding that satisfies the outsider's "participation constraint", but they would still keep the bulk of the joint surplus. Thus, broadly speaking, the different "bargaining" processes reconcile our result with Ornelas (2008).

Proposition 3 also indicates that the globally negotiated tariff binding is the farsighted MFN tariff $t^{fs}_{MFN}$. Moreover, the prospect of future FTA formation affects the farsighted MFN tariff when $b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$ but, as indicated in Proposition 3, jumps from $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$ to $t^{fs}_{MFN} = t^{pe}$ when $b \in \left[\bar{b}_{BND}, \bar{b}_{\emptyset}\right)$. Two sets of implications follow from this result; one set pertaining to the possibility of FTA formation itself and one set pertaining to the likelihood of subsequent FTA formation. While we recognize the stylized nature of our model (i.e. three symmetric countries), we believe this simple model provides some new insights that may factor into the complex evolution of international trade negotiations.

To focus on the first set of implications (i.e. those stemming from the possibility of FTA formation), suppose FTA formation will certainly take place so that $p = 1$. The first implication is that the shadow of future regionalism has a positive effect on the success of multilateral negotiations: multilateral tariff complementarity pushes the farsighted MFN tariff $t^{fs}_{MFN} = \frac{2}{3}t^{pe}$ below the politically efficient tariff $t^{pe}_{ij}(\emptyset) = t^{pe}$. Thus, governments' anticipation of future FTA formation, and their understanding that they would prefer lower global tariffs upon FTA formation, leads governments to incorporate multilateral tariff complementarity into the globally negotiated tariff bindings.

The second implication concerns the conditions governing the equilibrium emergence of binding overhang and tariff complementarity. When $b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$, global tariff negotiations in the shadow of FTA formation yield significant tariff concessions via

relatively low tariff bindings and to the extent that, in equilibrium, there is no binding overhang nor any individual tariff complementarity upon FTA formation. As discussed by Nicita et al. (2013), one could plausibly view the 1994 Uruguay Round of global tariff negotiations as essentially taking place between a small number of (relatively similar) advanced economies including the EU, US and Japan. Indeed, Beshkar et al. (2012) document that these countries had no binding overhang on 95–99% of HS 6-digit tariff lines in 2007. Moreover, recent cross-country empirical evidence from Gawande et al. (2012, 2015) estimates the US, Japan and major EU countries have some of the lowest values of $b$ in the world. In turn, given these countries have formed many FTAs, they have (essentially) not lowered tariffs on non-members and, thus, a lack of tariff complementarity has accompanied their FTAs. These observations are consistent with the predictions of our model when $b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$.

Conversely, when $b > \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$, global negotiations in the shadow of FTA formation yield relatively shallow tariff binding concessions and to the extent that, in equilibrium, FTA members practice binding overhang. In contrast to the empirical results of Limão (2006) and Karacaovali and Limão (2008) who find lower preferential tariff tariffs are not associated with lower external tariffs for the EU and US, Estevadeordal et al. (2008) find empirical evidence of tariff complementarity for South American FTA members. The former is consistent with our no binding overhang and no tariff complementarity results for countries with low $b$. Moreover, given Gawande et al. (2012, 2015) estimate that South American countries have substantially higher values of $b$ than the US, Japan and EU and Beshkar et al. (2012) document South American countries have substantial binding overhang, the latter is consistent with our binding overhang and tariff complementarity results for countries with high $b$. Thus, our theoretical results can reconcile the seemingly conflicting results of Limão (2006) and Karacaovali and Limão (2008) versus Estevadeordal et al. (2008).

The third implication concerns the mechanisms underlying the equilibrium emergence of binding overhang. For $\frac{1}{8}\varphi < b < \min\left\{\bar{b}_{\emptyset}, \bar{b}_{BND}\right\}$, the lack of binding overhang, and hence individual tariff complementarity, derives purely from the farsighted nature of globally negotiated tariff bindings. That is, "myopic" countries would negotiate a global tariff binding of $t^{pe}$ but, given $b > \frac{1}{8}\varphi$, tariff complementarity would then arise because $t^{*}_{IN} < t^{pe}$ when $b > \frac{1}{8}\varphi$. To this extent, the farsightedness of countries engaging in global tariff negotiations that take place in the shadow of subsequent FTA negotiations can help explain the lack of binding overhang in countries who were central figures in the 1994 Uruguay Round of negotiations such as the EU, US and Japan.

The mechanism underlying the equilibrium emergence of binding overhang fundamentally differs from that in Beshkar et al. (2015). There, the mechanism fundamentally revolves around uncertainty over political economy pressures and, in turn, cannot explain binding overhang with our notion of fixed political economy motivations (see their Proposition 1 on p. 5). For us, the fixed level of political economy pressures determine the common binding and, in the presence of FTAs, whether FTA members practice binding overhang. To empirically investigate the Beshkar et al. (2015) mechanism from this perspective, one could compare countries with the same number of FTAs and mean political economy pressure (and other relevant controls, including market power) but different volatility of political economy pressure. One would expect higher volatility leads to higher binding overhang. To empirically investigate our mechanism, one could compare countries with the same volatility of political economy pressure (and other relevant controls, including market power) but different mean levels of political economy volatility. One would expect a higher mean leads to larger increases in binding overhang upon FTA formation. Naturally, measuring time-varying political pressure presents a key challenge to such investigations.

The fourth implication concerns the effect of FTAs on trade flows. As discussed by Bergstrand et al. (2014, p. 3), changes in trade flows following FTAs are often used to infer the welfare effects of FTAs. Given our result regarding the absence of individual tariff complementarity, using FTA induced trade flow changes would seem to suggest that non-members suffer from FTAs. Similarly, given Ornelas (2008) finds world welfare rises upon an FTA if and only if one allows the insider to lower its external tariffs, FTA formation would appear to harm world welfare. However, this emphasizes the important point that, even though individual tariff complementarity does not arise upon FTA formation, its effect is embedded into the global tariffs *prior* to FTA formation. As such, our results suggest any effect of increased trade flows upon FTA formation due to individual tariff complementarity may already be embedded in the trade flows prior to the FTA. Thus, our results suggest that, via the farsighted nature of global tariff negotiations, the effect of an FTA on trade flows consists not only of the effect after the FTA forms but also the effect that the possibility of such an FTA taking place has on applied tariffs *prior to* FTA formation.

The second set of implications emerge from investigating the effect of changes in the likelihood of subsequent FTA formation. First, the farsighted MFN tariff $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$ is decreasing in $p$. That is, the shadow of future regionalism has a greater effect on global tariff negotiations when governments view future FTA formation as more likely because, in this case, governments care more about the impact of multilateral tariff complementarity whereby FTA formation lowers the jointly optimal tariff bindings.

Second, the extent to which our gated globalization result of a single FTA emerges in equilibrium (compared to a no-FTA equilibrium) depends on the likelihood of future FTA negotiations. Variation in $p$ does not affect the incentive of two countries to form a single FTA when presented with the opportunity; in Fig. 2, the $\underline{t}(b)$ curve is independent of $p$. However, $p$ does affect the political sacrifice governments are willing to suffer in order to prevent FTA formation. As $p$ falls, the farsighted MFN tariff $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$ moves closer to the politically efficient tariff $t^{pe}$. In turn, governments become less willing to sacrifice political efficiency in order to prevent FTAs which shifts the $\underline{t}(b) + x(b)$ curve down in Fig. 2. Thus, as Fig. 2 shows, the threshold $\bar{b}_{\emptyset}$ rises meaning that stronger political economy motivations are now required to prevent FTA formation. In this sense, FTAs are more likely to emerge when governments view FTA negotiations as less likely because applied tariffs in a world of gated globalization are closer to the politically efficient tariff $t^{pe}$.

## 6. Extensions

### 6.1. Asymmetric protocol

So far, we assumed the FTA formation protocol treats all country pairs symmetrically. While a fully-fledged analysis allowing arbitrary asymmetric protocols lies outside the scope of this paper, we can demonstrate robustness of our results to a completely random protocol. Section 6.1.1 does so using an asymmetric protocol where one country pair has a slightly higher probability than other country pairs of having the first FTA formation opportunity.

With an asymmetric protocol and governments conducting global negotiations by maximizing their joint payoff, an individual country's payoff could fall below that in the absence of global negotiations. Hence, we split Stage 0 of global negotiations into two sub-stages. Stage 0(a) mirrors Stage 0 from the baseline analysis: governments set the global tariff binding that maximizes their joint payoff. However, each country can now veto these global tariff bindings in Stage 0(b). Doing so generates a failed round of global negotiations, devoid of any tariff bindings, and leads directly

to FTA negotiations in Stage 1.[27] This endogenizes whether global negotiations precede FTA negotiations.

Our modified modeling of global negotiations also allows relaxing the property of our baseline analysis that countries cannot use the FTA formation process to influence concessions obtained during global negotiations. Section 6.1.2 explicitly investigates whether, *during* global negotiations, countries can extract larger concessions by threatening to veto global negotiations and let FTA formation take place instead. This again shows robustness of our main results and allows characterization of extracted concessions.

### 6.1.1. Robustness of main results

We now demonstrate robustness of our two main results by relaxing the assumption of a completely random protocol governing FTA negotiations. Specifically, we assume the country sequences $\{ij, ik, jk\}$ and $\{ij, jk, ik\}$ have probability $\frac{1}{6} + \frac{1}{2}\epsilon$ and the remaining country sequences have probability $\frac{1}{6} - \frac{1}{4}\epsilon$. That is, countries $i$ and $j$ have probability $\frac{1}{3} + \epsilon$ of being the first country pair that can form an FTA; the other two country pairs have the smaller probability $\frac{1}{3} - \frac{1}{2}\epsilon$. We refer to this protocol as the *Asymmetric Protocol of FTA Formation*. For the remainder of the current subsection, we assume $\epsilon > 0$ but small.

Two observations explain why global negotiations still prevent global free trade. First, in the absence of global negotiations, FTA formation still expands to global free trade when $b < \bar{b}_{OUT}$. After all, the FTA formation process begins after realization of the protocol ordering and, hence, its outcome does not depend on the probabilities associated with particular sequences of country pairs in the protocol. Second, the farsighted MFN tariff followed by a single FTA when FTA negotiations take place yields a higher joint expected government payoff than *any* tariff vector, including asymmetric tariff vectors, that leads to global free trade when FTA negotiations take place. Thus, allowing asymmetric tariff bindings does not alter our main result that global tariff negotiations prevent global free trade.[28]

While a single FTA can still emerge in equilibrium, like our baseline analysis, the asymmetric protocol complicates the underlying intuition by raising the issue of asymmetric tariff bindings. Given our robustness objective, we allow the tariff binding for the symmetric countries $i$ and $j$ to differ arbitrarily from country $k$'s tariff binding.[29] We refer to this possibility as "asymmetric tariff bindings", providing countries extra leverage to prevent FTAs.

Conditional on FTA expansion that yields global free trade, countries cannot exploit this extra leverage. Since country $k$'s tariff binding does not impact the attractiveness to countries $i$ and $j$ of forming the *sole* FTA, similar logic to our baseline analysis implies these countries always have an incentive to form an FTA in the absence of any other FTAs. However, asymmetric bindings imply that spokes may not form the last FTA that yields global free trade. In particular, the high binding spoke country may refuse FTA formation with the low binding spoke country and, anticipating the discrimination faced as spokes, each prospective spoke country may refuse any FTA formation with the prospective hub country. Such asymmetric tariff bindings would thus prevent FTA formation.

Two important considerations govern whether FTAs emerge after global negotiations. First, preventing FTA formation via asymmetric tariff bindings sacrifices political efficiency because a symmetric tariff binding maximizes governments' joint expected payoff conditional on any pattern of FTAs. Second, having prevented all FTAs, asymmetric tariff bindings are quite costly for a low binding country who now faces higher foreign tariffs than those faced by a high binding country. Indeed, any asymmetric tariff bindings that could prevent FTA formation are sufficiently costly for a low binding country that it will, as long as FTA negotiations are sufficiently likely, veto global negotiations and let FTAs expand to global free trade uninhibited by global tariff bindings. Thus, with sufficiently likely FTA negotiations, asymmetric tariff bindings are not optimal. Proposition 4 summarizes our discussion.

**Proposition 4.** *Consider the Asymmetric Protocol of FTA Formation where $\epsilon > 0$ but small. Further, suppose governments can impose asymmetric tariff bindings and can veto global negotiations in Stage 0(b). When $b < \bar{b}_{OUT}$, (i) global negotiations prevent global free trade and (ii) there exists a threshold $\bar{p} < 1$ such that, conditional on FTA negotiations, FTA formation takes place but does not lead to global free trade when $p > \bar{p}$.*

### 6.1.2. FTA formation and extraction of global tariff concessions

We now investigate whether, *during* global negotiations, countries can use the outcome of the FTA formation process to extract larger concessions. Formally, Stage 0(b) allows this possibility. To simplify the analysis while still elucidating the key economic forces, we initially suppose one country pair has the first FTA formation opportunity with certainty and countries impose a common tariff binding.[30] This is a special case of our Asymmetric Protocol of FTA Formation with $\epsilon = \frac{2}{3}$. We then generalize our results so that $\epsilon \in \left(0, \frac{2}{3}\right]$.

Global tariff bindings both increase and redistribute the joint expected payoff relative to global free trade: the payoff of the prospective insiders rise while that of the prospective outsider falls. Thus, when a country anticipates being an outsider with certainty and the associated discrimination under the global tariff bindings $t^{fs}_{MFN}$, it will veto global negotiations when FTA negotiations are sufficiently likely so that FTA formation expands uninhibited by global tariff bindings. Facing failed global negotiations, the prospective insiders agree to tariff binding concessions that mitigate the discrimination faced by the outsider. These tariff binding concessions come via a lower tariff binding when the insiders and the outsider are bound but come via a higher tariff binding when only the outsider is bound.[31] Nevertheless, these tariff binding concessions are sufficiently small that, as in our baseline analysis, the insiders still refuse subsequent FTA formation. Proposition 5 summarizes our results.

[27] Our approach here is formally equivalent to allowing FTA negotiations in a new Stage 0(c), rather than Stage 1, if a country vetoes global negotiations in Stage 0(b).

[28] In practice, developing countries gave far less than full reciprocity during the early GATT rounds (Hoda, 2001). On one hand, this could mitigate the extent that global negotiations prevent global free trade by increasing the attractiveness of FTAs in reducing high tariff bindings of developing countries. But, as discussed above in a *symmetric* setting, global tariff bindings create and redistribute surplus towards FTA members in a way that makes further FTA formation unattractive. In practice, this redistributive role of global negotiations would strengthen given the strongly asymmetric influence of developed countries over negotiations (Hoda, 2001). Indeed, in this context, exemptions from full reciprocity could be seen as a way to satisfy a developing country participation constraint.

[29] In practice, one could imagine that countries rule out tariff binding asymmetry that is vastly disproportional to the degree of asymmetry in country characteristics.

[30] If we allowed asymmetric tariff bindings between the insiders and the outsider, one of two things would happen. First, the endogenous choice of tariff bindings would render the "participation" constraint of the prospective FTA outsider, which is at the heart of our ensuing discussion, irrelevant. Second, the endogenous choice of tariff binding would somewhat tighten the participation constraint but ensuring it held would still require tariff binding concessions of the form we will discuss. Further, note that failure of global negotiations yields global free trade when $b < \bar{b}_{OUT}$, but Proposition 5 characterizes a symmetric tariff binding that yields a higher joint payoff and satisfies all relevant participation constraints. Thus, *enlarging* the class of tariff bindings to include asymmetric tariff bindings cannot generate failure of global negotiations.

[31] Ex-post, as in our baseline analysis, the insiders and the outsider would like to unilaterally raise their external tariff above the tariff binding when bound by the tariff binding. However, WTO rules permit the "withdrawal of equivalent concessions". Bagwell and Staiger (1999a) interpret this principle in terms of imports volumes. In our model, this implies a symmetric retaliatory tariff increase that ultimately leaves all countries worse off.

**Proposition 5.** *Let $b < \bar{b}_{OUT}$. Consider the Asymmetric Protocol of FTA Formation and suppose countries can veto global negotiations in Stage 0(b). If $\epsilon = \frac{2}{3}$ and $p < 1$ then*

(i) *Global negotiations prevent global free trade.*
(ii) *When FTA negotiations take place, a single FTA emerges between the pair of countries with the first FTA formation opportunity and the global tariff binding is $\hat{t}^{fs}_{MFN}$. If $\hat{t}^{fs}_{MFN}$ binds the external tariffs of the insiders and the outsider, $\hat{t}^{fs}_{MFN} = bd\left(1-\frac{p}{3}\right) + \Delta(b,p,\epsilon)$ and $\Delta(b,p,\epsilon) \leq 0$. If $\hat{t}^{fs}_{MFN}$ only binds the external tariff of the outsider, $\hat{t}^{fs}_{MFN} = bd + \Delta(b,p,\epsilon)$ and $\Delta(b,p,\epsilon) \geq 0$. In any case, $|\Delta(b,p,\epsilon)|$ is weakly increasing in $p$.*
(iii) *If $\epsilon < \frac{2}{3}$ then, for any $p$, (i) and (ii) hold and $|\Delta(b,p,\epsilon)|$ is weakly increasing in $\epsilon$.*

Proposition 5 again establishes the robustness of our main results, but also characterizes the concessions $\Delta(b,p,\epsilon)$ extracted by the prospective outsider. These concessions are weakly increasing in $p$: increases in $p$ strengthen the attractiveness to the prospective outsider of vetoing global negotiations in favor of FTA formation that, given $b < \bar{b}_{OUT}$, leads to global free trade. The reason for initially imposing $p < 1$ when $\epsilon = \frac{2}{3}$ is that $p = 1$ and $\epsilon = \frac{2}{3}$ imply one country actually becomes the outsider with certainty and thus extracts concessions yielding its global free trade payoff which requires a zero tariff binding. Nevertheless, while we introduced $p$ as the probability of FTA negotiations taking place, $p$ also has a discount-factor-like interpretation. Of course, given a network $g$ and a tariff binding vector $\bar{\tau}$,

$$\arg\max_{\bar{\tau}}(1-p)\sum_i G_i(\emptyset;\bar{\tau}) + p\sum_i G_i(g;\bar{\tau}) = \arg\max_{\bar{\tau}} \sum_i G_i(\emptyset;\bar{\tau}) + \frac{p}{1-p}\sum_i G_i(g;\bar{\tau}).$$

That is, $\bar{\tau}$ maximizes an immediate myopic payoff $\sum_i G_i(\emptyset;\bar{\tau})$ plus a discounted forward looking payoff $\frac{p}{1-p}\sum_i G_i(g;\bar{\tau})$ where $p \in (0,1)$ acts like a discount factor. Thus, $p < 1$ can be interpreted as governments anticipating that future FTA negotiations certainly take place but they place some weight on long term relative to short term outcomes. As such, the restriction of $p < 1$ when $\epsilon = \frac{2}{3}$ in Proposition 5 is not overly restrictive.

Proposition 5 emphasizes that the underlying motives for global tariff bindings in our baseline analysis do not include insurance motives. Even when the outsider's identity is known with certainty (i.e. $\epsilon = \frac{2}{3}$), the global tariff binding of Proposition 5 only differs from our baseline analysis via the participation constrains induced by veto power. That is, these global tariff bindings again represent the global efficiency implications of multilateral tariff complementarity and unilateral terms of trade incentives.

Proposition 5 (iii) establishes these results also apply when $\epsilon < \frac{2}{3}$ (for any $p \leq 1$). Intuitively, conditional on any single FTA outcome in the absence of veto power, the optimal global tariff binding is $t^{fs}_{MFN}$ and, in turn, $\epsilon$ merely alters the distribution over the identity of the prospective insiders without impacting the global tariff binding. However, Proposition 5 (iii) also characterizes the impact of uncertainty on global tariff bindings when countries have veto power. As $\epsilon$ rises, uncertainty over the prospective outsider's identity falls and this "likely" outsider extracts larger concessions: a lower global tariff binding when insiders and the outsider are bound, but a higher global tariff binding when only the outsider is bound. Intuitively, because global tariffs bindings redistribute surplus away from the eventual outsider and towards the eventual insiders, the "likely" outsider's veto threat over global negotiations becomes stronger. And, once global negotiations reduce its expected payoff below that resulting from FTA formation without any initial tariff bindings, this veto threat binds. In turn, the insiders grant tariff concessions and these concessions grow with $\epsilon$.

### 6.2. Exogenous global tariff bindings

Until now, our analysis endogenized the global tariff bindings. In our baseline analysis, the fundamental motivations for these bindings were the global efficiency implications of multilateral tariff complementarity and unilateral terms of trade incentives. Our extension in Section 6.1.2 allowed the "likely" outsider to extract concessions from the "likely" insiders, constituting the third fundamental motivation. However, naturally, our model ignores other real world motivations for global tariff bindings. Thus, one may wonder whether our results are robust to entering the FTA formation game with exogenous global tariff bindings.

We now ignore Stage 0 from our baseline analysis and assume FTA negotiations in Stage 1 take place with an exogenous common tariff binding $\bar{t}$. The key insight from our earlier analysis was that, through the extraction of non-trivial concessions from the eventual outsider, global negotiations reduce the eventual insiders' incentives to engage in subsequent FTA formation. Moreover, our analysis of the FTA formation game in the absence of global negotiations, where global free trade emerges for $b < \bar{b}_{OUT}$, is formally equivalent to imposing an exogenous common tariff binding $\bar{t}$ exceeding the Nash tariff $t_{Nash}$. Thus, Proposition 6 shows our main results hold if and only if the exogenous tariff binding $\bar{t}$ is tight enough.

**Proposition 6.** *Consider our baseline protocol for FTA formation where countries face a common exogenous tariff binding $\bar{t}$ at the beginning of Stage 1. Let $b < \bar{b}_{OUT}$. Then, there exists a threshold binding $\tilde{t}(b)$ such that (i) $t^{fs}_{MFN} < t^{*}_{IN} < \tilde{t}(b) < t^{*}_{OUT}$ and (ii) FTA formation leads to a single FTA when $\bar{t} < \tilde{t}(b)$ but leads to global free trade when $\bar{t} > \tilde{t}(b)$.*

Proposition 6 illustrates that the key insight from our baseline analysis is quite robust to various motivations driving global tariff bindings. When $b < \bar{b}_{OUT}$ in the absence of any tariff bindings, our earlier analysis showed FTA formation leads to global free trade. Here, the tariff barriers of the outsider are sufficiently high that the insiders use FTA formation with the outsider to reduce these barriers and political economy motivations are weak enough for the outsider that it participates in FTA expansion to global free trade. But, *any* global tariff binding that produces a world without binding overhang (and, by implication, no tariff complementarity), i.e. $\bar{t} < t^{*}_{IN}$, is one that prevents FTA formation. Here, the tariff binding severely constrains the outsider's tariff and delivers enough concessions to the insiders that they refuse FTA formation with the outsider. Further, for *any* global tariff binding $\bar{t} \in \left(t^{*}_{IN}, \tilde{t}(b)\right)$ a single FTA again emerges but tariff complementarity generates binding overhang for insiders. Thus, our key insight that concessions inherent in global tariff bindings can prevent global free trade is quite robust to exogenous tariff bindings.

### 6.3. Role of Article XXIV constraint on internal tariffs

Consistent with prior literature, we have imposed the GATT Article XXIV constraint that FTA members impose zero tariffs on each other, i.e. zero "internal" tariffs. This is natural given our interest in how global negotiations affect the extent of FTA formation in a world with WTO/GATT rules. But, while this requirement does not affect our key qualitative results, our analysis can shed light on some conceptual implications of Article XXIV.

To this end, we expand Stage 2 so that, along with the other tariff choices made in our baseline analysis, FTA members choose their internal tariffs to maximize their joint payoff. Naturally, internal tariffs are subject to any tariff bindings and a non-negativity

constraint. An immediate observation is that an FTA member's internal and external tariffs exhibit a form of tariff complementarity: a lower external tariff brings a lower internal tariff. A well known result in the literature is that joint determination of a common external tariff by Custom Union members helps preserve their preferential tariff margin that would otherwise be eroded by the traditional individual tariff complementarity effect. Our analysis shows that a common internal tariff plays a similar role: when external tariffs fall and lower the preferential margin of FTA members, internal tariffs also fall and partly restore this margin.

Our key results remain unaltered because the internal tariff non-negativity constraint binds, delivering zero internal tariffs, for $b < \frac{1}{8}\varphi$. Thus, for $b < \frac{1}{8}\varphi$, Propositions 2 and 3 apply.[32]

However, Article XXIV's zero internal tariff constraint bites once $b > \frac{1}{8}\varphi$. The politically efficient outcome in the presence of an FTA is for external *and* internal tariffs to equal $t^{pe} = bd$. Indeed, the individually optimal external tariff of insiders is also $t^{pe}$. But, to increase their preferential margin, insiders set internal tariffs below $t^{pe}$. Thus, from a political efficiency view, FTA internal tariffs are too low. In turn, given the complementarity between internal and external tariffs, governments would like to force insiders to set external tariffs above their individually optimal level. Two sources of political inefficiency thus arise: (i) like the baseline analysis, insiders lower their internal tariff below $t^{pe}$ and (ii) unlike the baseline analysis, the endogenous adjustment of FTA internal tariffs imply that governments cannot bind FTA external tariffs using the applied tariff that would maximize their joint payoff.

To mitigate these political inefficiencies, governments may set the tariff binding so low that the non-negativity constraint binds internal tariffs at zero. Indeed, there exists a range $b \in \left[\frac{1}{8}\varphi, \tilde{b}_{TC}\right]$ where $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$ achieves this objective. Further, when $b$ is sufficiently small, i.e. $b \in \left[\frac{1}{8}\varphi, \tilde{b}_{BND}\right]$ where $\tilde{b}_{BND} \in \left(\frac{1}{8}, \tilde{b}_{TC}\right)$, internal tariffs are far enough from $t^{pe} = bd$ that governments strategically set the tariff binding at $t^{fs}_{MFN}$ which endogenously constrains internal tariffs at zero. But, once $b$ gets sufficiently high, i.e. $b > \tilde{b}_{BND}$, internal tariffs are close enough to $t^{pe} = bd$ that governments jointly gain from relaxing the zero internal tariff constraint. Moreover, unlike the baseline analysis, governments can never prevent FTA formation when $b < b^*$. When governments strategically constrain internal tariffs at zero, this follows from the baseline analysis. And, otherwise, FTA members always benefit from coordinating optimal internal tariffs. Proposition 7 summarizes our discussion.

**Proposition 7.** *Suppose FTA members can set weakly positive internal tariffs. (i) For $b < \bar{b}_{OUT}$, global negotiations prevent global free trade. (ii) For $b \in \left[\bar{b}_{OUT}, b^*\right)$, a single FTA emerges in equilibrium when FTA negotiations take place and the global tariff binding is*

$$\tilde{t}^{fs}_{MFN} = \begin{cases} t^{pe}\left(1 - \frac{p}{3}\right) & \text{if } b < \tilde{b}_{BND} \\ t^{pe} & \text{if } b \geq \tilde{b}_{BND} \end{cases}$$

*where $\tilde{b}_{BND} \in \left(\frac{1}{8}\varphi, b^*\right)$ and $\tilde{b}_{BND} < \bar{b}_{BND}$. (iii) External tariffs are always $t^{fs}_{MFN}$ and FTA internal tariffs are strictly positive if and only if $b \geq \tilde{b}_{BND}$.*

Our baseline analysis and Proposition 7 shed light on some conceptual implications of Article XXIV. First, when $b \in \left[\bar{b}_{\emptyset}, b^*\right)$, *removal* of Article XXIV makes FTA formation feasible when it was otherwise infeasible. Indeed, in this range, external tariffs are capped at $t^{pe}$ in the absence of Article XXIV but exceed $t^{pe}$ in its presence. That is, removal of Article XXIV not only lowers internal tariffs of FTA members but, via endogenous adjustment of tariff bindings in the presence of complementarity between internal and external tariffs, also lowers the external tariffs of insiders and the outsider. Ultimately, relative to the absence of Article XXIV, Article XXIV can stymie FTA formation and global tariff liberalization via the stringent requirement of zero internal tariffs.

Nevertheless, second, Article XXIV can deliver lower global tariffs when a single FTA emerges irrespective of Article XXIV. When $b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$, Article XXIV delivers external tariffs of $t^{pe}\left(1 - \frac{p}{3}\right) < t^{pe}$ and zero internal tariffs while the absence of Article XXIV delivers external tariffs of $t^{pe}$ and strictly positive internal tariffs when $b > \tilde{b}_{BND}$. Here, removing Article XXIV not only raises internal tariffs of FTA members but, via endogenous adjustment of tariff bindings in the presence of complementarity between internal and external tariffs, also raises external tariffs of the insiders and the outsider. Presumably, these higher global tariffs underlie the fear rationalizing Article XXIV.

Yet, third, he presence of Article XXIV may have no effect on FTA formation or global tariffs. When $b < \min\left\{\bar{b}_{BND}, \bar{b}_{\emptyset}\right\}$ and $b < \tilde{b}_{BND}$ then, *irrespective* of Article XXIV, a single FTA emerges with external tariffs of $t^{pe}\left(1 - \frac{p}{3}\right) < t^{pe}$ and zero internal tariffs. Here, while the absence of Article XXIV allows positive internal tariffs, governments strategically set tariff bindings that deliver zero internal tariffs to avoid the political inefficiencies associated with low internal tariffs. That is, the presumed negative effects of removing Article XXIV's zero tariff constraint may not emerge in equilibrium because, in equilibrium, the extent of FTA formation and global tariffs can be independent of Article XXIV.

### 6.4. Incorporating exporter lobbying

Until now, political economy pressures emanated from the import competing sector. Indeed, given our effective partial equilibrium framework, individually optimal tariffs are independent of the parameter, say $\beta$, capturing additional weight that a government places on export sector producer surplus. But, in general, decisions over reciprocal tariff reductions (e.g. FTAs or globally negotiated tariff bindings) depend on $\beta$. Thus, we now allow $\beta > 0$.

Two reinforcing observations suggest the impact of exporter political pressure should not overturn our main result. First, absent global negotiations, exporter political pressure relaxes the political constraints inhibiting FTA expansion and makes global free trade more likely than our baseline analysis.[33] Second, our main result in the presence of global negotiations revolves around the idea that lower global tariff bindings relax the incentive for insiders to reduce the outsider's tariffs via subsequent FTA formation. With $\beta > 0$, the analogous expressions for our baseline tariff bindings of $t^{fs}_{MFN} = bd\left(1 - \frac{p}{3}\right)$ or $t^{fs}_{MFN} = bd$ are $(bd - \beta e)\left(1 - \frac{p}{3}\right)$ or $bd - \beta e$. Thus, our model says (i) positive bindings arise when $\beta < b\frac{d}{e}$ and (ii) bindings fall with $\beta$. In turn, given positive tariff bindings, insiders still refuse subsequent FTA formation that leads to global free trade.[34] Hence, these insights suggest global negotiations yielding positive tariff bindings still prevent global free trade with governments politically motivated by exporter and import-competing influences.

[32] Remember $\bar{b}_{OUT} < \frac{1}{8}\varphi$.

[33] In our baseline analysis that implicitly assumed $\beta = 0$: (i) $G_i(g^{FT}) > G_i(g_{jk})$ iff $b < \bar{b}_{OUT} = \frac{13}{137}\varphi$, (ii) $G_i(g^{FT}) > G_i(g_{ij})$ iff $b < \bar{b}_{IN} = \frac{101}{313}\varphi$ and (iii) $G_i(g_{ij}) > G_i(\emptyset)$ iff $b < \bar{b}_{FTA} = \frac{47}{299}\varphi$. When $\beta > 0$: (i) $G_i(g^{FT}) > G_i(g_{jk})$ iff $b < \bar{b}_{OUT}(\beta) = \frac{13}{137}\varphi + \beta\frac{176}{137}\frac{e}{d}$, (ii) $G_i(g^{FT}) > G_i(g_{ij})$ iff $b < \bar{b}_{IN}(\beta) = \frac{101}{313}\varphi + \beta\frac{616}{313}\frac{e}{d}$ and (iii) $G_i(g_{ij}) > G_i(\emptyset)$ iff $b < \bar{b}_{FTA}(\beta) = \frac{47}{299}\varphi + \beta\frac{440}{299}\frac{e}{d}$.

[34] That is, $G_i(g_{ij}) - G_i(g^{FT}) > 0$ always holds given $\beta < b\frac{d}{e}$ and, hence, the positive tariff bindings described in the text.

### 6.5. Global negotiations and alternative normative criteria

Implicitly, we follow the typical normative criterion in the literature for evaluating the relative merits of various liberalization processes: the possible attainment of global free trade. However, we could adopt other normative evaluation criteria.

The joint government payoff is one alternative criterion. This may seem odd when interpreting the reduced form parameter $b$ through a Grossman and Helpman (1994) framework of interest group influence. But, Baldwin (1987) argues that many other distributional concerns could microfound $b$; for example, $b$ could capture government concern over import-competing sector employment. Absent more targeted domestic instruments, tariffs may be the key instrument governments can manipulate. In turn, the joint government payoff could be a plausible normative criterion. Indeed, from this view, global negotiations always help by raising the joint government payoff above that under global free trade (Lemma 3).

World welfare represents a second alternative criterion. If FTA negotiations take place with certainty and lead to global free trade then, from an ex-post view, global negotiations reduce world welfare. However, our modeling of the uncertainty over FTA negotiations taking place allows an alternative ex-ante view. Indeed, even when global negotiations prevent global free trade, global negotiations still increase ex-ante expected world welfare when FTA negotiations are sufficiently unlikely.[35] Here, the ability to negotiate reciprocal tariff reductions from their relatively high non-cooperative level outweighs the unlikely possibility that FTAs will completely rid the world of these non-cooperative tariffs.[36] Again, this alternative normative criterion paints global negotiations in a positive light.

## 7. Conclusion

Multilateralism can influence regionalism in many ways. An important channel is via the impact of globally negotiated tariff bindings on incentives for subsequent FTA formation. Indeed, the key question in our paper, the effect of global tariff negotiations on FTA formation, addresses an important gap in the literature. In their survey, Freund and Ornelas (2010, p. 156) note that there is a "… scarcity of analyses on how multilateralism affects regionalism". Our analysis is a first step in this direction. While our symmetric competing exporters model is highly stylized, none of our results rely on the knife edge case of symmetry and, therefore, introducing some moderate exogenous asymmetry into our model will leave our results qualitatively unaffected. Moreover, our results offer insights that are more general than the competing exporters model because they rely on economic forces that should be present independent of the underlying trade model.

First, given multilateralism and FTAs coexist and represent alternative pathways to global free trade, our main result is that multilateralism via global tariff negotiations can actually cause a world stuck short of global free trade. The basic economic intuition here is twofold. First, in a world where FTAs represent the only path towards global free trade, FTA formation represents an attractive way to reduce the high non-cooperative tariffs that would prevail in the absence of FTAs. Indeed, unless governments have sufficiently strong political economy motivations, this can propel FTA formation to global free trade. Second, by reducing tariffs worldwide, multilateralism mitigates the need for countries to use FTAs as a means to lower the tariffs of their trading partners. As such, multilateralism can be the reason FTA formation stops short of global free trade. This twofold logic is more general than the stylized setup of the competing exporters model.

Second, our result that a fragmented world of gated globalization with a single FTA can emerge highlights a tension dating back to at least Bagwell and Staiger (2005b). In a general economic environment, they show the politically efficient tariff in the absence of FTAs, $t^{pe}$, could be vulnerable to reciprocal bilateral tariff reductions. However, our forward looking model highlights that countries may set tariff bindings different from $t^{pe}$ in order to deter subsequent FTAs. Of course, whether countries do so depends on how much political efficiency would be sacrificed. Moreover, if FTA formation can be deterred, it will be due to strong political economy motivations of governments which not only requires a sufficiently large $b$ but also a sufficiently high tariff binding because this makes the import competing sector strong and, thus, valuable to protect. Again, the logic underlying our gated globalization result is not specific to the competing exporters model.

Third, while our result that the shadow of regionalism affects multilateral negotiations rests on the concept of multilateral tariff complementarity, this concept was first identified by Ornelas (2008) in a more general economic environment than ours. Moreover, in contrast to Ornelas where multilateral tariff complementarity takes place after FTA formation, our results highlight that forward looking countries build multilateral tariff complementarity into global tariff negotiations prior to FTA formation taking place. Thus, multilateral tariff complementarity may play an important role in shaping global tariff bindings even though it will not be observed in practice following FTA formation. In turn, the common practice of using observations regarding tariff complementarity or changes in trade flows upon FTA formation for inferring welfare changes may require re-examination.

Our results can help shed light on conflicting empirical stylized facts in the literature regarding binding overhang and individual tariff complementarity. For the major participants in global tariff negotiations such as the EU, US and Japan, our model rationalizes the observed absence of binding overhang and individual tariff complementarity in countries with relatively low political economy motivations. Yet, for countries with relatively high political economy motivations such as various South American countries, our model also rationalizes the observed presence of binding overhang and individual tariff complementarity.

Future research could ask how the FTA formation process resulting from one round of global negotiations may itself affect the outcome of subsequent rounds (that may be partly anticipated earlier). Doing so would have to recognize differences in the qualitative nature and the role played by global negotiations before and after FTA formation. Specifically, FTA formation creates a fragmented world riddled with discrimination and, thus, global negotiations must deal with various participation constraints and outside options (especially when not allowing direct transfers). Our approach developed here could provide a basis for handling this issue.

Additionally, given the 1994 Uruguay Round of negotiations covered bound tariffs of all WTO members, not only the few key negotiating (and developed) countries, one could extend our analysis to negotiations between highly asymmetric countries. An interesting possibility worthy of exploration is whether such a model could deliver asymmetries in the FTA formation incentives of developing and developed countries.

## Appendix A. Welfare expressions

The individual components of welfare can be expressed for an arbitrary vector of global tariffs $\tau$: $CS_i = \frac{1}{18}\left(2e + d - \sum_{j\neq i} t_{ij}\right)^2 +$

[35] Formally, we can show this is true when $p < \bar{p}(b)$ where $\lim_{b\to 0} \bar{p}(b) = 1$, $\lim_{b\to \frac{1}{3}\varphi} \bar{p}(b) > 0$ and $\bar{p}(b)$ is continuously decreasing in $b$.

[36] In cases where FTA negotiations would not lead to global free trade in the absence of global negotiations, global negotiations always yield higher expected world welfare than an FTA formation process that, with probability $p$, takes place in the absence of global negotiations.

$\frac{1}{18}\sum_{j\neq i,k\neq i,j}(2e+d+2t_{ji}-t_{jk})^2$, $PS_i^I = \frac{d}{3}\left[3\alpha-(2e+d)+\sum_{j\neq i}t_{ij}\right]$, $PS_i^Z = \frac{e}{3}\left[3\alpha-(2e+d)+t_{zj}-2t_{zi}\right]$ for $Z \neq I$ and $z \neq i \neq j$ and $TR_i = \frac{1}{3}\sum_{j\neq i,k\neq i,j}t_{ij}\,(e-d+t_{ik}-2t_{ij})$.

## Appendix B. Proofs

Before presenting the proofs, we address two notation issues. The first issue relates to government payoffs. Specifically, $G_i(g)$ and $G(g)$ denote the respective payoffs received by the government of country $i$ and the joint government payoff given a network of FTAs $g$ with the possible networks described in Section 3.2.

The second issue relates to tariffs. We let $t_{IN}$ and $t_{OUT}$ denote arbitrary applied tariffs of, respectively, the insiders and outsider with $t^*_{IN}$ (see Eq. (9)) and $t^*_{OUT} \equiv t_{Nash}$ (see Eq. (8)) denoting the respective optimal applied tariffs. Moreover, as described in Section 3, $\tau$ denotes the vector of tariffs. But, we let (i) $\tau(t)$ denote a tariff vector where all countries impose a common tariff $t$ (i.e. $t_{ij} = t$ for all $i,j$), (ii) $\tau_{-ij}(t)$ denote the vector $\tau(t)$ *except that* countries $i$ and $j$ set zero tariffs on each other, and (iii) $\tau^{FTA}_{-ij}(t)$ denote the vector that (potentially) differs from $\tau_{-ij}(t)$ because $t_{ik} = t_{jk} = \min\{t^*_{IN}, t\}$ and $t_{ki} = t_{kj} = \min\{t^*_{OUT}, t\}$.

We now present three lemmas that will be used in the proofs of lemmas and propositions from the main text.

**Lemma 5.** *Suppose $G_i(g^{FT}) > G_i\left(g_j^H\right)$. Then, $g^{FT}$ is the equilibrium outcome of the FTA formation game if (i) $G_i(g^{FT}) > max\{G_i(g_{jk}), G_i(g_{ij})\}$ and (ii) $G_i(g_{ij}) > G_i(\emptyset)$.*

**Proof.** Stage 1(c): $g = g_j^H$ for some country $j$ at the beginning of Stage 1(c). Symmetry and $G_i(g^{FT}) > G_i\left(g_j^H\right)$ implies $a_i = a_k = J$ and thus $g^{FT}$ emerges in Stage 1(c).

Stage 1(b): $g = g_{ij}$ for some countries $i$ and $j$ at the beginning of Stage 1(b). Given Stage 1(c) and symmetry, $G_i(g^{FT}) > max\{G_i(g_{jk}), G_i(g_{ij})\}$ implies $a_h = J$ for each country $h$ in the last active pair. Thus, an FTA forms in Stage 1(b).

Stage 1(a): $g = \emptyset$ at the beginning of Stage 1(a). Given Stages 1(b)-(c) and symmetry, $G_i(g^{FT}) > G_i(g_{ij}) > G_i(\emptyset)$ implies $a_h = J$ for each country $h$ in the last active pair. Hence, an FTA forms in Stage 1(a) and $g^{FT}$ is the equilibrium outcome. ■

**Lemma 6.** *Suppose $G_i(g^{FT}) > G_i\left(g_j^H\right)$. Then, for some countries $i$ and $j$, $g_{ij}$ is the unique equilibrium outcome of the FTA formation game if (i) $G_i(g^{FT}) < max\{G_i(g_{jk}), G_i(g_{ij})\}$ and (ii) $G_i(g_{ij}) > G_i(\emptyset)$. The first active pair form this FTA if $G_i(g_{ij}) > G_i(g_{jk})$ but the last active pair forms the FTA if $G_i(g_{ij}) < G_i(g_{jk})$.*

**Proof.** Stage 1(c): $g = g_j^H$ for some country $j$ at the beginning of Stage 1(c). Symmetry and $G_i(g^{FT}) > G_i\left(g_j^H\right)$ implies $a_i = a_k = J$ and thus $g^{FT}$ emerges in Stage 1(c).

Stage 1(b): $g = g_{ij}$ for some countries $i$ and $j$ at the beginning of Stage 1(b). But, given Stage 1(c) and symmetry, $G_i(g^{FT}) < max\{G_i(g_{jk}), G_i(g_{ij})\}$ implies $a_h = NJ$ for some country $h$ in each active pair. Thus, $g_{ij}$ remains and Stage 1(c) is never attained.

Stage 1(a): $g = \emptyset$ at the beginning of Stage 1(a). Given Stages 1(b)-(c) and symmetry, $G_i(g_{ij}) > G_i(\emptyset)$ implies $a_h = J$ for each country $h$ in the last active pair. Thus, an FTA forms in Stage 1(a). In turn, given the sequential protocol, $g_{ij}$ is the unique equilibrium outcome for some countries $i$ and $j$. If $G_i(g_{ij}) < G_i(g_{jk})$, then $a_h = NJ$ for some country $h$ in the first two active pairs. Thus, the last active pair form the FTA. Conversely, if $G_i(g_{ij}) > G_i(g_{jk})$ then $a_h = J$ for each country $h$ in the second active pair and, in turn, for each country in the first active pair. Thus, the first active pair form the FTA. ■

**Lemma 7.** *Suppose $G_i(g^{FT}) > G_i\left(g_j^H\right)$. Then, $\emptyset$ is the equilibrium outcome of the FTA formation game if $G_i(\emptyset) > G_i(g_{ij})$ and either (i) $G(\emptyset) > G(g^{FT})$ or (ii) $G_i(g^{FT}) < max\{G_i(g_{jk}), G_i(g_{ij})\}$.*

**Proof.** Note, $g_j^H$ cannot emerge in equilibrium because symmetry and $G_i(g^{FT}) > G_i\left(g_j^H\right)$ imply $a_h = J$ for each spoke country $h = i,k$ in Stage 1(c). There are now two cases to consider.

First, let $G(\emptyset) > G(g^{FT})$. Then, $G_i(\emptyset) > \max\{G_i(g^{FT}), G_i(g_{ij})\}$ given symmetry and $G_i(\emptyset) > G_i(g_{ij})$. In turn, each country $h$ of an active pair in Stage 1(a) chooses $a_h = NJ$. Hence, no FTAs form. Second, let $G_i(g^{FT}) < max\{G_i(g_{jk}), G_i(g_{ij})\}$. This implies $a_h = NJ$ for some player $h$ in any active pair in Stage 1(b) and, hence, $g_{ij}$ remains after Stage 1(b) and Stage 1(c) is never attained. In turn, $G_i(\emptyset) > G_i(g_{ij})$ implies $a_h = NJ$ for each country $h$ in any active pair in Stage 1(a) and no FTAs form. ■

We now move on to proofs of propositions and lemmas from the main text.

**Proof of Lemma 1.** In Stage 1(c) of the FTA formation game, $G_i(g^{FT}) > G_i\left(g_j^H\right)$ if and only if $b < \frac{1}{3}\varphi + \frac{7}{6d}t_K$ where $t_K$ is the common tariff of spokes. This must hold given Eqs. (3) and (5) say that non-prohibitive tariffs require $b < \frac{1}{3}\varphi$ (see Section 3.3.1). Further, $t_K = t^*_{IN}$ in the absence of global negotiations (see Eq. (10)) and $t_K = \min\{t^*_{IN}, t\}$ where $t$ is the global tariff binding in the presence of global negotiations.

**Proof of Lemma 2.** Assume a single FTA emerges conditional on FTA negotiations taking place. First, suppose the tariff bindings $\tau$ bind the applied tariffs of insiders and, given $t^*_{IN} < t^*_{OUT} = t_{Nash}$, the outsider. Then, Eq. (16) implies the optimal tariff bindings are $\tau\left(t^{pe}\left(1-\frac{p}{3}\right)\right)$. Note, $\tau\left(t^{pe}\left(1-\frac{p}{3}\right)\right)$ binds the applied tariffs of the insiders and the outsider if and only if $t^{pe}\left(1-\frac{p}{3}\right) \leq \min\{t^*_{IN}, t^*_{OUT}\} = t^*_{IN}$ which reduces to

$$b \leq \bar{b}_{TC} \equiv \frac{3}{24-11p}\varphi. \tag{19}$$

Second, suppose the tariff bindings $\tau$ do not bind insiders' applied tariffs. Then, Eqs. (13)–(14) say the optimal tariff bindings are $\tau(t^{pe})$. Eqs. (5), (8), (9) and (13) imply these tariff bindings bind the applied tariffs of insiders, i.e. $t^{pe} < t^*_{IN}$, if and only if $b < \frac{1}{8}\varphi$ and of the outsider, i.e. $t^{pe} < t^*_{OUT}$, for any non-prohibitive tariff. Thus, let $b \geq \frac{1}{8}\varphi$ hereafter.

The optimal tariff bindings are now determined by comparing governments' joint expected payoff under these two cases. Note that, for $b \geq \frac{1}{8}\varphi$,

$$\begin{aligned}&\left[pG\left(g_{ij};\tau_{-ij}\left(t^{pe}\left(1-\frac{p}{3}\right)\right)\right)+(1-p)\,G\left(\emptyset;\tau\left(t^{pe}\left(1-\frac{p}{3}\right)\right)\right)\right]\\&-\left[pG\left(g_{ij};\tau^{FTA}_{-ij}\left(t^{pe}\right)\right)+(1-p)\,G\left(\emptyset;\tau^{FTA}_{-ij}\left(t^{pe}\right)\right)\right]\\&=\frac{1}{1089}p\left[b^2d^2\left(121p-144\right)-30bd\left(e-d\right)+6(e-d)^2\right]\end{aligned} \tag{20}$$

with Eq. (20) being positive if and only if $b < \bar{b}_{BND}$ where

$$\bar{b}_{BND} \equiv \frac{11\sqrt{9-6p}-15}{144-121p}\varphi \tag{21}$$

with $\bar{b}_{BND} \in \left(\frac{1}{8}\varphi, \bar{b}_{TC}\right)$ for $p \in (0,1]$. Finally, we verify that $t^*_{IN} > t^{pe}\left(1-\frac{p}{3}\right)$ for $b < \bar{b}_{BND}$ and $t^*_{IN} \leq t^{pe} \leq t^*_{OUT}$ for $b \geq \bar{b}_{BND}$ noting that $\bar{b}_{BND} \geq \frac{1}{8}\varphi$ and $p \in (0,1]$. First, $t^*_{IN} > t^{pe}\left(1-\frac{p}{3}\right)$ for $b < \bar{b}_{BND}$ follows

because $\bar{b}_{TC} > \bar{b}_{BND}$ given one can verify that $z(p) \equiv \bar{b}_{TC} - \bar{b}_{BND}$ is increasing in $p$ and $z(0) = 0$. Second, (i) $t^*_{IN} \leq t^{pe}$ reduces to $b \geq \frac{1}{8}\varphi$, which holds given $b \geq \bar{b}_{BND} > \frac{1}{8}\varphi$, and (ii) $t^*_{OUT} > t^{pe}$ holds for any $b < \frac{1}{3}\varphi$. Thus, $\bar{b}_{BND} < \bar{b}_{TC}$ implies $t^{fs}_{MFN}$ binds external tariffs except that $t_{ik}(g_{ij}) = t^*_{IN} < t^{fs}_{MFN} = bd$ when $b \geq \bar{b}_{BND}$. ■

**Proof of Lemma 3. (i).** Suppose a single FTA has formed. Given Lemma 1, a hub-spoke network cannot emerge in equilibrium. Thus, subsequent FTA formation yields $g^{FT}$. We now show $\Delta_1 \equiv G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{fs}_{MFN}\right)\right) - G_i\left(g^{FT}\right) > 0$ and hence, conditional on $g_{ij}$, $a_h = NJ$ for some insider $h = i,j$ in Stage 1(b) of the FTA formation game meaning the outcome at the end of Stage 1(b) remains a single FTA and Stage 1(c) is never attained. When $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$, $\Delta_1 > 0$ for $b < \frac{2}{3(1-p)}\varphi$ and hence $\Delta_1 > 0$ for any $b < \frac{1}{3}\varphi$. When $t^{fs}_{MFN} = t^{pe}$, $\frac{\partial \Delta_1}{\partial b} \leq 0$ for $b \leq \frac{52}{229}\varphi$ and $\frac{\partial \Delta_1}{\partial b} \geq 0$ for $b \geq \frac{52}{229}\varphi$. Hence, $\Delta_1$ is minimized for $b = \frac{52}{229}\varphi$ in which case $\Delta_1 = \frac{7}{687}(e-d)^2 > 0$.

**(ii).** Given $g_{ij}$ does not expand further, will two countries form an FTA? Given Lemma 1 and the previous paragraph, Lemmas 6–7 (see beginning of Appendix B) say yes if and only if $G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{fs}_{MFN}\right)\right) - G_i\left(\emptyset; \tau\left(t^{fs}_{MFN}\right)\right) > 0$ which represents an "insider participation constraint" (IPC). The general form of the IPC for tariff bindings $\tau(t)$ is

$$f\left(t^*_{IN}, t^*_{OUT}, t_{Nash}, t\right) \equiv G_i\left(g_{ij}; \tau^{FTA}_{-ij}(t)\right) - G_i\left(\emptyset; \tau\left(\min\left\{t, t_{Nash}\right\}\right)\right) > 0. \quad (22)$$

When $t < t^*_{IN}$ then $t_{IN} = t_{OUT} = t$ and $f(\cdot) > 0$ reduces to $t < \frac{2}{3}(e-d) - 2bd \equiv \underline{t}_1(b)$. Moreover, letting $b < \bar{b}_{BND}$ and $t \equiv t^{fs}_{MFN}$, we have $t = t^{pe}\left(1 - \frac{p}{3}\right) < t^*_{IN}$ given $\bar{b}_{BND} < \bar{b}_{TC}$. In turn, $t = t^{fs}_{MFN} < \underline{t}_1(b)$ reduces to $b < \frac{2}{(9-p)}\varphi$ which holds given $\bar{b}_{BND} < \frac{2}{(9-p)}\varphi$. Thus, a single FTA emerges when $b < \bar{b}_{BND}$.

**(iii).** When $b < \bar{b}_{BND}$ a single FTA emerges when FTA negotiations occur and, by definition, $t^{fs}_{MFN}$ maximizes governments' joint expected payoff. When $b \geq \bar{b}_{BND}$, $t^{fs}_{MFN} = t^{pe}$ and, if FTA negotiations occur, either a single FTA or no FTA emerges. Again, if a single FTA emerges, $t^{fs}_{MFN}$ maximizes governments' joint expected payoff by definition. Moreover, the joint government payoff is $G(\emptyset; \tau(t^{pe}))$ if no FTAs emerge which is, in fact, the highest joint payoff governments can achieve; in particular, $G(\emptyset; \tau(t^{pe})) > G\left(g^{FT}\right)$. ■

**Proof of Proposition 1.** Suppose global tariff negotiations take place. Then, Proposition 3 states that a single FTA emerges in equilibrium when $b < \bar{b}_{\emptyset}$. Moreover, the proof of Proposition 3 establishes that no FTAs emerge in equilibrium when $b \geq \bar{b}_{\emptyset}$. ■

**Proof of Proposition 2.** In the presence of global tariff negotiations, Proposition 1 implies $g^{FT}$ is not the equilibrium outcome of the FTA formation game. However, in the absence of global tariff negotiations, Lemma 5 (see beginning of Appendix B) implies $g^{FT}$ is the equilibrium outcome of the FTA formation game when $b < \bar{b}_{OUT}$. Given Lemma 1, the conditions of Lemma 5 hold for $b < \bar{b}_{OUT}$ because, using the expressions in Appendix A, we have $\bar{b}_{OUT} < \bar{b}_{FTA} < \bar{b}_{IN}$ where (i) $G_i\left(g^{FT}\right) - G_i(g_{ij}) > 0 \iff b < \bar{b}_{IN} \equiv \frac{101}{313}\varphi$ and (ii) $G_i(g_{ij}) - G_i(\emptyset) > 0 \iff b < \bar{b}_{FTA} \equiv \frac{47}{299}\varphi$. ■

**Proof of Lemma 4.** Given Lemma 1, Lemmas 5–6 (see beginning of Appendix B) imply $G_i(g_{ij}) > G_i(\emptyset)$ is a sufficient condition for equilibrium FTA formation. Thus, $G_i(g_{ij}) \leq G_i(\emptyset)$ is a necessary condition for preventing FTA formation. For tariff bindings $\tau(t)$, the general form for failure of this "insider participation constraint" (IPC) is

$$f\left(t^*_{IN}, t^*_{OUT}, t_{Nash}, t\right) \equiv G_i\left(g_{ij}; \tau^{FTA}_{-ij}(t)\right) - G_i\left(\emptyset; \tau\left(\min\left\{t, t_{Nash}\right\}\right)\right) \leq 0. \quad (23)$$

Two cases establish that a necessary condition for $f(\cdot) \leq 0$ is that $t$ exceed a threshold $\underline{t}(b)$. First, let $t < t^*_{IN}$. Then, as described in the proof of Lemma 3, $t_{IN} = t_{OUT} = t$ and $f(\cdot) > 0$ reduces to $t < \frac{2}{3}(e-d) - 2bd \equiv \underline{t}_1(b)$. Second, let $t \in \left[t^*_{IN}, t^*_{OUT}\right)$. Then, $t_{IN} = t^*_{IN}$ and $t_{OUT} = t$. $f(\cdot) > 0$ now reduces to $t \notin (\underline{t}_2(b), \bar{t}_2(b))$ where $\underline{t}_2(b) \equiv \hat{t}(b) - \frac{3}{77}v(\theta)^{1/2}$ and $\bar{t}_2(b) \equiv \hat{t}(b) + \frac{3}{77}v(\theta)^{1/2}$ with $\hat{t}(b) \equiv \frac{e-d}{7} + \frac{6}{7}bd$ and $v(\theta) \equiv [50bd + 13(e-d)][8bd - (e-d)]$. Thus, noting that $t^*_{OUT} > \hat{t}(b)$ for any $b < \frac{1}{3}\varphi$, a necessary condition for $f(\cdot) \leq 0$ is $t \geq \underline{t}(b)$ where

$$\underline{t}(b) \equiv \begin{cases} \underline{t}_1(b) = \frac{2}{3}(e-d) - 2bd & \text{if } t < t^*_{IN} \\ \underline{t}_2(b) = \hat{t}(b) - \frac{3}{77}v(\theta)^{1/2} & \text{if } t \geq t^*_{IN} \end{cases}. \quad (24)$$

We now show that $f(\cdot) > 0$ when $b < \frac{1}{8}\varphi$. First, let $t < t^*_{IN}$. Then, $t < \underline{t}_1(b)$ because $\underline{t}_1(b) > t^*_{IN}$ reduces to $b < \frac{19}{75}\varphi$. Thus, $f(\cdot) > 0$ if $b < \frac{1}{8}\varphi$. Second, let $\left[t^*_{IN}, t^*_{OUT}\right)$. Then, $f(\cdot) \leq 0$ if and only if $t \in \left[\underline{t}_2(b), \bar{t}_2(b)\right]$. But, this interval exists if and only if $v(\theta) \geq 0$ which reduces to $b \geq \frac{1}{8}\varphi$. Thus, $f(\cdot) > 0$ if $b < \frac{1}{8}\varphi$. Finally, let $t \geq t^*_{OUT}$. Then, $f(\cdot) > 0$ reduces to $b < \bar{b}_{FTA}$ where the proof of Proposition 2 gives $\frac{1}{8}\varphi < \bar{b}_{FTA} \equiv \frac{47}{299}\varphi$. Thus, $f(\cdot) > 0$ if $b < \frac{1}{8}\varphi$. ■

**Proof of Proposition 3.** To begin, note that we use Lemmas 5–7 introduced at the beginning of Appendix B as well as the definition of $\underline{t}(b)$ from the proof of Lemma 4.

Define $b^*$ such that $t^{pe}(b) \geq \underline{t}(b)$ iff $b \geq b^*$. Thus, $b^* \approx .177\varphi > \frac{1}{8}\varphi$. By definition of $t^{pe}$, $G(\emptyset; \tau(t^{pe})) \geq G(g; \tau)$ for any network of FTAs $g$ and any tariff bindings $\tau$. Thus, when $b \geq b^*$, Lemma 7 implies no FTAs emerge for the tariff bindings $\tau(t^{pe})$. In turn, $\tau(t^{pe})$ are the optimal tariff bindings when $b \geq b^*$. Thus, hereafter, we let $b < b^*$. In turn, $t^{pe}(b) < \underline{t}(b)$ and, by definition of $\underline{t}(b)$, $G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{pe}\right)\right) > G_i(\emptyset; \tau(t^{pe}))$ hereafter.

We now establish that a single FTA emerges in equilibrium when FTA negotiations take place and the tariff bindings are $\tau\left(t^{fs}_{MFN}\right)$ as described in Lemma 2. Lemma 3 established this when $b < \bar{b}_{BND}$. Thus, we now let $b \geq \bar{b}_{BND}$ and verify the two conditions needed for Lemma 6. Note that $b \geq \bar{b}_{BND}$ implies $t^{fs}_{MFN} = t^{pe} > t^*_{IN}$. Thus, first, as noted above, $G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{pe}\right)\right) > G_i(\emptyset; \tau(t^{pe}))$ given $b < b^*$. Second, the proof of Lemma 3 established $\Delta_1 = G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{pe}\right)\right) - G_i\left(g^{FT}\right) > 0$.

By construction, $\tau\left(t^{fs}_{MFN}\right)$ maximizes the expected joint government payoff conditional on $g_{ij}$ emerging when FTA negotiations take place; in particular, governments achieve a higher joint expected payoff than by choosing $\tau(0)$ which corresponds with global free trade. Further, Lemma 1 rules out an equilibrium hub-spoke network. Thus, the only possible equilibrium outcome apart from $g_{ij}$ is $\emptyset$.

Lemmas 5 and 6 imply $G_i(\emptyset) \geq G_i(g_{ij})$ is a necessary condition for no FTAs in equilibrium. However, noting that $b^* < \frac{19}{75}\varphi$, the proof of Lemma 4 established that $G_i(g_{ij}) > G_i(\emptyset)$ when (i) $b < \frac{1}{8}\varphi$ and (ii) $b \in \left[\frac{1}{8}\varphi, b^*\right)$ and the tariff bindings are $\tau(t)$ where $t < t^*_{IN}$. Thus, we hereafter consider $b \in \left[\frac{1}{8}\varphi, b^*\right)$ and $t \geq t^*_{IN}$. We now see that a single FTA emerges if and only if $b < \bar{b}_{\emptyset}$ noting that $x(b)$ emerges from solving

$$G(\emptyset; \tau(t)) - \left[p \cdot G\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{fs}_{MFN}\right)\right) + (1-p) \cdot G\left(\emptyset; \tau\left(t^{fs}_{MFN}\right)\right)\right] \geq 0. \quad (25)$$

Specifically, Eq. (25) reduces to $t \in [t^{pe} - x(b), t^{pe} + x(b)]$ where

$$x(b) = \begin{cases} \frac{1}{3}bd\left(-p^2 + 6p\right)^{1/2} > 0 & \text{if } b < \bar{b}_{BND} \\ \frac{(6p)^{1/2}}{33}\left[bd\left(97bd - 5(e-d)\right) + (e-d)^2\right]^{1/2} > 0 & \text{if } b \geq \bar{b}_{BND} \end{cases}. \quad (26)$$

Let $b < \bar{b}_{\emptyset}$ noting that $z(b) \equiv t^{pe} + x(b) - \underline{t}(b)$ is strictly increasing in $b$ with $z\left(\bar{b}_{\emptyset}\right) = 0$. Then, $t^{pe} + x(b) < \underline{t}(b)$ and, in turn, there is no $\tau(t)$ such that $G_i(\emptyset) \geq G_i(g_{ij})$ and Eq. (25) holds. Hence, the optimal tariff bindings are $\tau\left(t^{fs}_{MFN}\right)$ as described in Lemma 2 and $g_{ij}$ is the equilibrium outcome. Lemma 2 implies $\tau\left(t^{fs}_{MFN}\right)$ binds all external tariffs except those of insiders when $b \in \left[\bar{b}_{BND}, \bar{b}_{\emptyset}\right)$ in which case $t_{IN} = t^{*}_{IN} < t^{pe}$.

Finally, let $b \geq \bar{b}_{\emptyset}$. Then, given $z(b)$ is strictly increasing in $b$, $t^{pe} + x(b) > \underline{t}(b)$. Thus, the tariff bindings $\tau(t)$ with $t = \underline{t}_2(b) > t^{pe}$ imply that $G_i(\emptyset) \geq G_i(g_{ij})$ and that Eq. (25) holds. Given Eq. (7) implies $G(\emptyset; \tau(t)) > G(g^{FT})$, Lemma 7 implies no FTAs emerge in equilibrium if the tariff bindings are $\tau(\underline{t}_2(b))$. In turn, given $G(\emptyset; \tau(t))$ is decreasing in $t$ for $t > t^{pe}$, $\tau\left(t^{fs}_{MFN}\right) = \tau(\underline{t}_2(b))$ are the optimal tariff bindings for $b \in [b_{\phi}, b^{*})$. The proof is complete upon recognizing that, by definition, $t^{pe} = \underline{t}_2(b)$ for $b = b^{*}$.■

**Proof of Proposition 4. (i).** For the game without Stage 0, the outcome of Stages 1-3 is independent of the realized protocol. Hence, given $b < \bar{b}_{OUT}$, global free trade emerges in equilibrium.

Now consider the game with Stage 0. For a symmetric tariff binding, the logic in the proof of Proposition 5 implies $\Delta(b, p, \epsilon) = 0$ when $\epsilon > 0$ but sufficiently small. Thus, the optimal symmetric tariff bindings are $\tau\left(t^{fs}_{MFN}\right)$ where $t^{fs}_{MFN} = t^{pe}\left(1 - \frac{p}{3}\right)$ and, conditional on FTA negotiations, a single FTA emerges. Hence, if asymmetric tariff bindings emerge in equilibrium then the joint expected government payoff exceeds $\mathrm{E}\left[G\left(g_{ij}; t^{fs}_{MFN}\right)\right] \equiv pG\left(g_{ij}; \tau^{FTA}_{-ij}\left(t^{fs}_{MFN}\right)\right) + (1-p)\, G\left(\emptyset; \tau\left(t^{fs}_{MFN}\right)\right)$. But, for an arbitrary tariff binding vector $\tau$, $\arg\max_{\tau} pG\left(g^{FT}\right) + (1-p)\, G(\emptyset; \tau) = \tau(bd)$. Yet, $\mathrm{E}\left[G\left(g_{ij}; t^{fs}_{MFN}\right)\right] - \left[pG\left(g^{FT}; \tau(bd)\right) + (1-p)\, G(\emptyset; \tau(bd))\right] = \frac{1}{9}(bd)^2 p(3+p) > 0$. Hence, $g^{FT}$ cannot emerge in equilibrium.

**(ii).** Given the optimal symmetric tariff binding above, FTA formation takes place in equilibrium (conditional on FTA negotiations) unless there exist asymmetric tariff bindings that prevent FTA formation. Thus, if asymmetric bindings are optimal then, given $\epsilon$ is sufficiently small, the joint government payoff exceeds $\mathrm{E}\left[G\left(g_{ij}; t^{fs}_{MFN}\right)\right]$. Moreover, given (i) a symmetric binding $t_h = t$ for $h = i, j$ and (ii) either $t < t^{*}_{IN}$ or $b < \frac{1}{8}\varphi$, then Eq. (24) is independent of country $k$'s binding and implies $G_h(g_{ij}; \tau) > G_h(\emptyset; \tau)$ for $h = i, j$. First, suppose $G_h\left(g^{FT}\right) > G_h\left(g^{H}_{h'}\right)$ for $h \neq h'$. Then, $\emptyset$ is not an equilibrium outcome because $i$ and $j$ benefit by forming an FTA which, by the logic of Lemmas 5–6, will either remain in place or expand to $g^{FT}$. Thus, if $\emptyset$ is an equilibrium outcome then $G_h\left(g^{H}_{h'}\right) > G_h\left(g^{FT}\right)$ for some $h \neq h'$ so that $g^{H}_{h'}$ remains in place if it emerges in Stage 1(b). Hence, second, suppose this is so and consider a tariff binding vector $\tau$ where an arbitrary country $z$ has the lowest tariff binding (countries $i$ and $j$ have a symmetric binding). Then, we can show numerically that there exists $\bar{p} < 1$ such that $p > \bar{p}$ implies $pG_z\left(g^{FT}\right) + (1-p)\, G_z(\emptyset; \tau(t_{Nash})) > G_z(\emptyset; \tau)$ when $G(\emptyset; \tau) > \mathrm{E}\left[G\left(g_{ij}; t^{fs}_{MFN}\right)\right]$. Hence, given $p > \bar{p}$, country $z$ vetoes global negotiations in Stage 0(b) whenever $G(\emptyset; \tau) > \mathrm{E}\left[G\left(g_{ij}; t^{fs}_{MFN}\right)\right]$. Thus, $\emptyset$ is not an equilibrium outcome when $p > \bar{p}$. ■

**Proof of Proposition 5.** The proof proceeds by establishing variants of Lemmas 1–4. In turn, for $b < \bar{b}_{OUT}$, the proofs of Propositions 1–3 apply again. Note that, given $G_i(g_{ij}) > G_i(g_{jk})$ in what follows, Lemma 6 implies the first country pair form the single FTA where $p_{ij}$ denotes the probability that $i$ and $j$ are the first country pair in the Asymmetric Protocol of FTA Formation.

Lemma 1 still applies given spokes impose symmetric tariffs because of a symmetric tariff binding or symmetric individually optimal tariffs.

For Lemma 2, Stage 0(b) introduces participation constraints. Let $\mathrm{E}[G_i(t)] \equiv p\left[\sum_{h,h'} p_{hh'} G_i\left(g_{hh'}; \tau^{FTA}_{-hh'}(t)\right)\right] + (1-p)\; G_i(\emptyset; \tau(t))$ and $\mathrm{E}[G(t)] \equiv \sum_i \mathrm{E}[G_i(t)]$. Further, let (i) $G^{*}_i \equiv \mathrm{E}\left[G_i\left(t^{fs}_{MFN}\right)\right]$ and $G^{*} \equiv \sum_i G^{*}_i$ and (ii) $\hat{G}_i \equiv pG_i\left(g^{FT}\right) + (1-p)\, G_i(\emptyset; \tau(t_{Nash})) \leq G_i\left(g^{FT}\right)$ and $\hat{G} = \sum_i \hat{G}_i \leq G\left(g^{FT}\right)$, with strict inequality if $p < 1$. Then, given $t = t^{fs}_{MFN}$ and $b < \bar{b}_{OUT}$, country $i$ vetoes global negotiations in Stage 0(b) if and only if $\hat{G}_i > G^{*}_i$. More generally, conditional on a single FTA, the global tariff binding maximizes $\mathrm{E}[G(t)]$ subject to $\mathrm{E}[G_h(t)] \geq \hat{G}_h$ for all $h$. Note that $G^{*}_h > \hat{G}_h$ always holds for the first pair of countries.

Let $t < t^{*}_{IN} < t^{*}_{OUT}$ and $t^{fs}_{MFN} \equiv bd\left(1 - \frac{p}{3}\right)$. Suppose $\hat{G}_h > G^{*}_h$ for the country not in the first country pair. Note that (i) $\mathrm{E}[G_h(t)]$ is strictly concave in $t$ if $p < 1$ or $\epsilon < \frac{2}{3}$, decreasing in $t$ if $p = 1$ and $\epsilon = \frac{2}{3}$, and $\frac{\partial \mathrm{E}[G_h(t)]}{\partial t}|_{t = t^{fs}_{MFN}} < 0$ and (ii) $\mathrm{E}[G(t)]$ is strictly concave in $t$ with $\frac{\partial \mathrm{E}[G(t)]}{\partial t}|_{t = t^{fs}_{MFN}} = 0$. Solving $\mathrm{E}[G_h(t)] = \hat{G}_h$ yields $t \equiv \hat{t}^{fs,1}_{MFN} = t^{fs}_{MFN} + \Delta(b, p, \epsilon) \geq 0$ where (i) the inequality is strict when $p < 1$ or $\epsilon < \frac{2}{3}$, (ii) $\Delta(b, p, \epsilon) \leq 0$, and (iii) given $\hat{G}_h$ is increasing in $p$ and independent of $\epsilon$ but $\mathrm{E}[G_h(t)]$ is decreasing in both $p$ and $\epsilon$ then $|\Delta(b, p, \epsilon)|$ is increasing in $p$ and $\epsilon$. Thus, $\mathrm{E}\left[G\left(\hat{t}^{fs,1}_{MFN}\right)\right] > \mathrm{E}[G(0)] = G\left(g^{FT}\right) \geq \hat{G}$ and $t = \hat{t}^{fs}_{MFN}$ maximizes $\mathrm{E}[G(t)]$ subject to $\mathrm{E}[G_h(t)] \geq \hat{G}_h$ for all $h$. Moreover, given $p < 1$ or $\epsilon < \frac{2}{3}$, $\mathrm{E}\left[G_h\left(\hat{t}^{fs,1}_{MFN}\right)\right] = \hat{G}_h \leq G_h\left(g^{FT}\right)$ implies $\mathrm{E}\left[G_i\left(\hat{t}^{fs,1}_{MFN}\right)\right] > \mathrm{E}[G_i(0)] = G_i\left(g^{FT}\right) \geq \hat{G}_i$. Thus $t = \hat{t}^{fs,1}_{MFN}$ is the optimal symmetric tariff binding for $t < t^{*}_{IN}$.

Now let $t^{*}_{IN} < t < t^{*}_{OUT}$ and $t^{fs}_{MFN} \equiv t^{pe}$. Also, redefine $\hat{t}^{fs,1}_{MFN}$ so that $\Delta(b, p, \epsilon) \equiv 0$ if $\hat{G}_h < G^{*}_h$. Suppose $\hat{G}_h > G^{*}_h$ where country $h$ is not in the first country pair. Unlike above, we now have $\frac{\partial \mathrm{E}[G_h(t)]}{\partial t}|_{t = t^{fs}_{MFN}} > 0$. Solving $\mathrm{E}[G_h(t)] = \hat{G}_h$ yields $t \equiv \hat{t}^{fs,2}_{MFN} = t^{fs}_{MFN} + \Delta(b, p, \epsilon) \geq 0$ where (i) the inequality is strict when $p < 1$ or $\epsilon < \frac{2}{3}$, (ii) $\Delta(b, p, \epsilon) \geq 0$, and (iii) given $\hat{G}_h$ is increasing in $p$ and independent of $\epsilon$ but $\mathrm{E}[G_h(t)]$ is decreasing in both $p$ and $\epsilon$ then $|\Delta(b, p, \epsilon)|$ is increasing in $p$ and $\epsilon$. A necessary condition for $\hat{t}^{fs,2}_{MFN}$ to be the optimal binding is $\mathrm{E}\left[G\left(\hat{t}^{fs,2}_{MFN}\right)\right] > \mathrm{E}\left[G\left(\hat{t}^{fs,1}_{MFN}\right)\right]$. Indeed, this is also a sufficient condition because then $\mathrm{E}\left[G\left(\hat{t}^{fs,2}_{MFN}\right)\right] > \mathrm{E}\left[G\left(\hat{t}^{fs,1}_{MFN}\right)\right] > G\left(g^{FT}\right) \geq \hat{G}$ which, given $\mathrm{E}\left[G_h\left(\hat{t}^{fs,2}_{MFN}\right)\right] = \hat{G}_h$, implies $\mathrm{E}\left[G_i\left(\hat{t}^{fs,2}_{MFN}\right)\right] > \mathrm{E}[G_i(0)] = G_i\left(g^{FT}\right) \geq \hat{G}_i$. Otherwise, $\hat{t}^{fs,1}_{MFN}$ is the optimal binding. Hereafter, let $\hat{t}^{fs}_{MFN}$ denote the optimal binding.

Lemma 4 applies again because it is independent of the equilibrium tariff bindings or the participation constraints introduced by Stage 0(b).

Lemma 3 again applies given three observations. First, $\hat{t}^{fs}_{MFN} < \underline{t}(b)$ because $b < \bar{b}_{OUT}$ implies $b < \frac{1}{8}\varphi$ and the proof of Lemma 4 establishes that $t \geq \underline{t}(b)$ cannot hold when $b < \frac{1}{8}\varphi$. Second, $\Delta_1 \equiv G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(\hat{t}^{fs}_{MFN}\right)\right) - G_i\left(g^{FT}\right) > 0$ given that (i) $\mathrm{E}\left[G_i\left(\hat{t}^{fs}_{MFN}\right)\right] > G_i\left(g^{FT}\right)$ requires $\max\left\{G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(\hat{t}^{fs}_{MFN}\right)\right), G_i\left(\emptyset; \tau\left(\hat{t}^{fs}_{MFN}\right)\right)\right\} > G_i\left(g^{FT}\right)$ and (ii) $\hat{t}^{fs}_{MFN} < \underline{t}(b)$ implies $G_i\left(g_{ij}; \tau^{FTA}_{-ij}\left(\hat{t}^{fs}_{MFN}\right)\right) > G_i\left(\emptyset; \tau\left(\hat{t}^{fs}_{MFN}\right)\right)$. Third, $\hat{t}^{fs}_{MFN} > 0$ implies $\mathrm{E}\left[G\left(\hat{t}^{fs}_{MFN}\right)\right] > \sum_i G_i\left(g^{FT}\right)$.

**Proof of Proposition 6.** Note that $G_i\left(g^{FT}\right) > G_i\left(g^{H}_{j}\right)$ for any common tariff imposed by the spokes and $\tilde{t}(b)$ is defined such that $G_i(g_{ij}) > G_i(g^{FT})$ if and only if $\bar{t} < \tilde{t}(b)$. Thus, for $\bar{t} < \tilde{t}(b)$, Lemmas 5–6 imply that the equilibrium outcome is a single FTA if $G_i(g_{ij}) > G_i(\emptyset)$ but no FTAs if $G_i(g_{ij}) > G_i(\emptyset)$.

We now establish that $t^{fs}_{MFN} < t^{*}_{IN} < \tilde{t}(b) < t^{*}_{OUT}$ for $b < \bar{b}_{OUT}$. First, $t^{fs}_{MFN} < t^{*}_{IN}$ follows from $\bar{b}_{OUT} < \bar{b}_{BND}$. Second, conditional on $\bar{t} < t^{*}_{IN}$, $G_i(g_{ij}) > G_i(g^{FT})$ if and only if $\bar{t} < \bar{t}_1(b)$. Moreover, (i) $\bar{t}_1(b) - t^{*}_{IN} \propto e - d + bd > 0$ and (ii) $\bar{t}_1(b) - t^{fs}_{MFN} > 0$ reduces to $b < \bar{b}(p) \equiv \frac{2}{3(1-p)}\varphi$ which always holds given $\bar{b}(p) \geq \frac{2}{3}\varphi > \bar{b}_{OUT}$. Third, conditional on $\bar{t} \in [t^{*}_{IN}, t^{*}_{OUT}]$, $G_i(g_{ij}) > G_i(g^{FT})$ if and only if $\bar{t} < \bar{t}_2(b)$. Moreover,

(i) $\bar{t}_2(b) - t^*_{IN} > 0$ for any $b > 0$ and (ii) $\bar{t}_2(b) - t^*_{OUT} < 0$ for any $b < \frac{101}{313}\varphi$ which holds for any $b < \bar{b}_{OUT}$. Fourth, conditional on $\bar{t} > t^*_{OUT}$, $G_i(g_{ij}) > G_i(g^{FT})$ if and only if $b > \frac{101}{313}\varphi$ which never holds for any $b < \bar{b}_{OUT}$. Thus, $t^{fs}_{MFN} < t^*_{IN} < \tilde{t}(b) < t^*_{OUT}$ for $b < \bar{b}_{OUT}$.

Finally, let $\bar{t} > \tilde{t}(b)$ so that $G_i(g^{FT}) > G_i(g_{ij})$. As outlined above, this implies $\bar{t} > t^*_{IN}$. Conditional on $\bar{t} > t^*_{IN}$, $G_k(g^{FT}) - G_k(g_{ij}) > 0$ if $b < \frac{13}{137}\varphi = \bar{b}_{OUT}$. Thus, Lemma 5 implies $g^{FT}$ is the equilibrium outcome if $G_i(g_{ij}) - G_i(\emptyset) > 0$. This is established in the proof of Lemma 4 for $b < \frac{1}{8}\varphi$ which must hold for $b < \bar{b}_{OUT}$. ■

**Proof of Proposition 7. (i).** The proof proceeds by showing that, for $b \leq \frac{1}{8}\varphi$ and hence $b < \bar{b}_{OUT}$, the individually optimal and jointly optimal tariffs are unchanged from the baseline analysis and hence the baseline analysis again applies. In particular, strictly positive internal tariffs do not emerge in equilibrium when $b \leq \frac{1}{8}\varphi$.

We begin by describing the optimal tariffs in the absence of tariff bindings. For $g = \emptyset$, the optimal external tariffs of country $i$ are $t_{ih}(g) = \frac{1}{4}(e - d + 3bd)$ for $h = j,k$. For $g = g_{ij}$, the optimal external and internal tariffs are given by

$$t_{ij}(g_{ij}; t_{ik}) = \frac{1}{7}\left[-(e-d) + 3bd + 5t_{ik}(g_{ij})\right] \tag{27}$$

$$t_{ik}(g_{ij}; t_{ij}) = \frac{1}{11}\left[(e-d) + 3bd + 7t_{ij}(g_{ij})\right] \tag{28}$$

$$t_{kh}(g_{ij}) = \frac{1}{4}\left[(e-d) + 3bd\right] \text{ for } k = i,j. \tag{29}$$

Solving Eq.s (27)–(28) by ignoring any non-negativity constraints, $t_{ik}(g_{ij}) = t^*_{EXT} = bd > 0$ and $t_{ij}(g_{ij}) = t^*_{INT} = \frac{1}{7}[-(e-d) + 8bd]$ where $t^*_{INT} > 0$ if and only if $b > \frac{1}{8}\varphi$. Further, in general, $t_{ij}(g_{ij}) > 0$ requires $t_{ik}(g_{ij}) > \underline{t}_{EXT} = \frac{1}{5}[(e-d) - 3bd]$ where $\underline{t}_{EXT} > 0$ if and only if $b < \frac{1}{3}\varphi$. For $g = g^H_i$ and $h = j,k$, the optimal external and internal tariffs are given by $t_{jk}(g^H_i) = t_{jk}(g_{ij})$, $t_{ji}(g^H_i) = t_{ji}(g_{ij})$ and $t_{ih}(g^H_i) = \frac{1}{2}[-(e-d) + 3bd]$ where $t_{ih}(g^H_i) > 0$ if and only if $b > \frac{1}{3}\varphi$. Letting $g^{FT}$ denote the network of three FTAs then, for $h = j,k$ we have $t_{ih}(g^{FT}) = t_{ih}(g^H_i)$. Given all second order conditions hold independently of any tariff levels, any internal tariff is zero when it violates the non-negativity constraint. With zero internal tariffs when $b \leq \frac{1}{8}\varphi$, the external tariffs mirror the baseline analysis. Hence, ignoring any tariff binding considerations, Propositions 1–3 again apply.

Indeed, noting that Eqs. (27)–(28) are independent of $t_{kh}(g_{ij})$ for $k \neq i,j$, the consideration of tariff bindings does not alter the property that $t_{ij}(g_{ij}) \equiv 0$ for any solution of Eqs. (27)–(28) subject to non-negative tariff constraints when $b \leq \frac{1}{8}\varphi$. Specifically, there is no tariff binding $t$ that binds the external tariff $t_{EXT}(t) \equiv t_{ik}(g_{ij}; t_{ij} = t_{INT}(t))$ and yields a strictly positive internal tariff $t_{INT}(t) \equiv t_{ij}(g_{ij}; t_{ik} = t) > 0$ when $b \leq \frac{1}{8}\varphi$. To see this note that, using Eqs. (27)–(28), $t_{EXT}(t) = \frac{1}{11}[6bd + 5t]$ so that $t \leq t_{EXT}(t)$ if and only if $t \leq \hat{t} \equiv bd$. But, $t \leq \hat{t}$ and $t_{INT}(t) > 0$ cannot both hold when $b \leq \frac{1}{8}\varphi$ because (i) $t_{INT}(t) > 0$ requires $t > \underline{t}_{EXT}$ yet (ii) $t \leq \underline{t}_{EXT}$ for all $t \leq \hat{t}$ when $b \leq \frac{1}{8}\varphi$ because $\hat{t} > \underline{t}_{EXT}$ requires $b > \frac{1}{8}\varphi$.

**(ii)–(iii).** Given we just established that the baseline results apply for $b \leq \frac{1}{8}\varphi$, we now proceed by establishing variants of Lemmas 1–4 for $b > \frac{1}{8}\varphi$ and, for a global tariff binding $t$, $t_{ij}(g_{ij}; t_{ik}, t) > 0$.

For Lemma 1 and global tariff bindings $\tau(t)$, $G_k(g^{FT}) > G_k(g^H_i)$ only if $b \lesssim .179\varphi$ regardless of whether $t$ binds $t_{ik}(g_{ij}; t_{ij}, t)$. Note that $b^* < .179\varphi$ and, thus, we hereafter restrict attention to $b \leq b^*$.

For Lemma 2, suppose the optimal binding $\tilde{t}^{fs}_{MFN}$ weakly binds the external tariffs of the insider and the outsider, which requires $\tilde{t}^{fs}_{MFN} \leq t^*_{EXT}$. Following the logic of Lemma 2, $\tilde{t}^{fs}_{MFN} = \frac{3p(e-d)+bd(147-23p)}{147-20p}$ with $\tilde{t}^{fs}_{MFN} = bd$ for $p = 0$ and $\frac{\partial \tilde{t}^{fs}_{MFN}}{\partial p} > 0$. Hence, $\tilde{t}^{fs}_{MFN} > t^*_{EXT}$ for all $p > 0$. Since the second order condition holds, the optimal binding $\tilde{t}^{fs}_{MFN}$ such that $\tilde{t}^{fs}_{MFN} \leq t^*_{EXT}$ is $\tilde{t}^{fs}_{MFN} = t^*_{EXT}$. Now suppose the optimal binding only binds the external tariff of the outsider so that $t_{ik}(g_{ij}; t_{ij}, t) < \tilde{t}^{fs}_{MFN} < t_{kh}(g_{ij})$. Then, the optimal binding is $\tilde{t}^{fs}_{MFN} = bd$. Note that setting a tariff binding $t < \underline{t}_{EXT}$ implies zero internal tariffs and, from the baseline analysis, an optimal tariff binding that binds the insider and the outsider of $\tilde{t}^{fs}_{MFN} = t^{fs}_{MFN} = bd\left(1 - \frac{p}{3}\right)$ as long as $t^{fs}_{MFN} < \underline{t}_{EXT}$ which reduces to $b < \tilde{b}_{TC} = \frac{3}{24-5p}\varphi \leq \bar{b}_{TC}$. Binding the insiders and the outsiders via $t^{fs}_{MFN}$ and ensuring zero internal tariffs yields a higher joint expected government payoff than only binding the outsider via $\tilde{t}^{fs}_{MFN} = bd$ and allowing positive internal tariffs if and only if $b < \tilde{b}_{BND} \equiv \frac{7(36-6p)^{1/2}-6}{288-49p}\varphi$ where, for $p > 0$, $\tilde{b}_{BND} < \tilde{b}_{TC}$ and $\tilde{b}_{BND} < \bar{b}_{BND}$.

Lemma 3 again applies given three observations. First, $f(\cdot) > 0$ for $t^{fs}_{MFN}$ follows from the baseline analysis given $\tilde{b}_{BND} < \bar{b}_{BND}$ and, otherwise, the analogy of $f(\cdot) > 0$ follows since $f(\cdot; t) \propto (6bd + (e-d) - 7t)^2$ for any tariff binding $t$ that only strictly binds the outsider. Second, $\Delta_1 > 0$ for $t^{fs}_{MFN}$ follows from the baseline analysis and $\Delta_1 > 0$ for $\tilde{t}^{fs}_{MFN} \neq t^{fs}_{MFN}$ follows because $\Delta_1 = 0$ has no real solution and $\Delta_1 > 0$ when $b = 0$. Third, $t^{fs}_{MFN} > 0$ and $\tilde{t}^{fs}_{MFN} > 0$ implies $\mathrm{E}[G(g_{ij}; t)] \equiv pG\left(g_{ij}; \tau^{FTA}_{-ij}(t)\right) + (1-p)G(\emptyset; \tau(t)) > \sum_i G_i(g^{FT})$ for $t = t^{fs}_{MFN}, \tilde{t}^{fs}_{MFN}$.

Lemma 4 can now be restated so that there is no tariff binding $t$ that prevents FTA formation when $b < b^*$. First, suppose $t \leq \underline{t}_{EXT}$ so there are zero FTA internal tariffs. Then, given the proof of Proposition 3, a necessary condition for preventing FTA formation is $t \geq bd$ when $b < b^*$. However, this contradicts the supposition that $t \leq \underline{t}_{EXT}$ given $b > \frac{1}{8}\varphi$ implies $bd > \underline{t}_{EXT}$. Second, suppose $t > \underline{t}_{EXT}$ so there are strictly positive FTA internal tariffs. Then, given the previous paragraph, FTA formation cannot be prevented because $f(\cdot; t) \propto ((e-d) + 6bd - 7t)^2 > 0$ if $t$ only binds the outsider and $f(\cdot; t) \propto ((e-d) - 3bd + 2t)^2 > 0$ if $t$ binds insiders and the outsider. ■

## References


Aghion, P., Antrās, P., Helpman, E., 2007. Negotiating free trade. J. Int. Econ. 73 (1), 1–30.

Amador, M., Bagwell, K., 2013. The theory of optimal delegation with an application to tariff caps. Econometrica 81 (4), 1541–1599.

Bagwell, K., Staiger, R., 1999a. An economic theory of GATT. Am. Econ. Rev. 89 (1), 215–248.

Bagwell, K., Staiger, R., 1999b. Regionalism and multilateral tariff cooperation. In: Pigott, J., Woodland, A. (Eds.), International Trade Policy and the Pacific Rim. Macmillan.,

Bagwell, K., Staiger, R., 2005a. Enforcement, private political pressure, and the General Agreement on Tariffs and Trade/World Trade Organization escape clause. J. Leg. Stud. 34 (2), 471–513.

Bagwell, K., Staiger, R., 2005b. Erratum to multilateral trade negotiations, bilateral opportunism and the rules of GATT/WTO. J. Int. Econ. 67 (2), 268–294.

Baldwin, R., 1987. Politically realistic objective functions and trade policy PROFs and tariffs. Econ. Lett. 24 (3), 287–290.

Bergstrand, J.H., Larch, M., Yotov, Y., 2014. Economic integration agreements, border effects, and distance elasticities in the gravity equation. (Mimeo).

Beshkar, M., Bond, E., Rho, Y., 2012. Tariff binding and overhang: theory and evidence. (Mimeo).

Beshkar, M., Bond, E., Rho, Y., 2015. Tariff binding and overhang: theory and evidence. J. Int. Econ. 97 (1), 1–13.

Bhagwati, J., 1991. The World Trading System at Risk. Princeton University Press.

Bhagwati, J., 1993. Regionalism and multilateralism: an overview. In: Panagriya, A., de Melo, J. (Eds.), New Dimensions in Regional Integration. World Bank.,

Estevadeordal, A., Freund, C., Ornelas, E., 2008. Does regionalism affect trade liberalization toward nonmembers? Q. J. Econ. 123 (4), 1531–1575.

Ethier, W.J., 1998. Regionalism in a multilateral world. J. Polit. Econ. 106 (6), 1214–1245.

Freund, C., 2000. Multilateralism and the endogenous formation of preferential trade agreements. J. Int. Econ. 52 (2), 359–376.

Freund, C., Ornelas, E., 2010. Regional trade agreements. Annu. Rev. Econ. 2 (1), 139–166.

Gawande, K., Bandyopadhyay, U., 2000. Is protection for sale? Evidence on the Grossman-Helpman theory of endogenous protection. Rev. Econ. Stat. 82, 139–152.

Gawande, K., Krishna, P., Olarreaga, M., 2012. Lobbying competition over trade policy. Int. Econ. Rev. 53 (1), 115–132.

Gawande, K., Krishna, P., Olarreaga, M., 2015. A political-economic account of global tariffs. Econ. Polit.

Goldberg, P.K., Maggi, G., 1999. Protection for sale: an empirical investigation. Am. Econ. Rev. 89, 1135–1155.
Grossman, G., Helpman, E., 1994. Protection for sale. Am. Econ. Rev. 84, 833–850.
Hoda, A., 2001. Tariff Negotiations and Renegotiations Under the GATT and the WTO: Procedures and Practices. Cambridge University Press.
Horn, H., Maggi, G., Staiger, R., 2010. Trade agreements as endogenously incomplete contracts. Am. Econ. Rev. 100 (1), 394–419.
Karacaovali, B., Limão, N., 2008. The clash of liberalizations: preferential vs. multilateral trade liberalization in the European Union. J. Int. Econ. 74 (2), 299–327.
Krishna, P., 1998. Regionalism and multilateralism: a political economy approach. Q. J. Econ. 113 (1), 227–251.
Lake, J., 2017. Free Trade Agreements as dynamic farsighted networks. Econ. Inq. 55 (1), 31–50.
Levy, P., 1997. A political-economic analysis of free-trade agreements. Am. Econ. Rev. 87 (4), 506–519.
Limão, N., 2006. Preferential trade agreements as stumbling blocks for multilateral trade liberalization: evidence for the United States. Am. Econ. Rev. 96 (3), 896–914.
Maggi, G., 1999. The role of multilateral institutions in international trade cooperation. Am. Econ. Rev. 89, 190–214.
Nicita, A., Olarreaga, M., Silva, P., 2013. Cooperation in WTO's tariff waters. (Mimeo).
Ornelas, E., 2005a. Endogenous free trade agreements and the multilateral trading system. J. Int. Econ. 67 (2), 471–497.
Ornelas, E., 2005b. Trade creating free trade areas and the undermining of multilateralism. Eur. Econ. Rev. 49 (7), 1717–1735.
Ornelas, E., 2008. Feasible multilateralism and the effects of regionalism. J. Int. Econ. 74 (1), 202–224.
Richardson, M., 1993. Endogenous protection and trade diversion. J. Int. Econ. 34 (3), 309–324.
Riezman, R., 1999. Can bilateral trade agreements help to induce free trade? Can. J. Econ. 32 (3), 751–766.
Saggi, K., Yildiz, H., 2010. Bilateralism, multilateralism, and the quest for global free trade. J. Int. Econ. 81 (1), 26–37.
Stoyanov, A., Yildiz, H., 2015. Preferential versus multilateral trade liberalization and the role of political economy. Eur. Econ. Rev. 80, 140–164.