ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Saggi, Kamal; Wong, Woan Foong; Yildiz, Halis Murat

Working Paper

The Role of Non-Discrimination in a World of Discriminatory Preferential Trade Agreements

CESifo Working Paper, No. 8139

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Saggi, Kamal; Wong, Woan Foong; Yildiz, Halis Murat (2020) : The Role of Non-Discrimination in a World of Discriminatory Preferential Trade Agreements, CESifo Working Paper, No. 8139, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/216535

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

CESifo WORKING PAPERS

8139
2020

February 2020

The Role of Non-Discrimination in a World of Discriminatory Preferential Trade Agreements

Kamal Saggi, Woan Foong Wong, Halis Murat Yildiz

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8139*

The Role of Non-Discrimination in a World of Discriminatory Preferential Trade Agreements

Abstract

In a three-country model of endogenous trade agreements, we study the implications of the Most Favored Nation Clause (MFN) when countries are free to form discriminatory preferential trade agreements (PTAs). While PTA members discriminate against non-member countries, MFN requires non-members to treat PTA members in a non-discriminatory fashion. We show that MFN reduces the potency of a country's optimal tariffs and therefore its incentive for unilaterally opting out of trade liberalization. Thus, MFN can be a catalyst for trade liberalization. However, when PTAs take the form of customs unions, the efficiency case for MFN as well as its pro-liberalization effect is weaker since one country finds itself *deliberately excluded* by member countries as opposed to staying out *voluntarily*.

JEL-Codes: F110, F120.

Keywords: trade agreements, tariffs, customs unions, World Trade Organization, coalition proof Nash equilibrium, welfare.

Kamal Saggi
Department of Economics
Vanderbilt University
USA – Nashville, TN 37235-1828
kamal.saggi@vanderbilt.edu

Woan Foong Wong
Department of Economics
University of Oregon
USA – Eugene, Oregon, 97403
wfwong@uoregon.edu

Halis Murat Yildiz
Department of Economics
Ryerson University
Canada – Toronto, ON M5B 2K3
hyildiz@ryerson.ca

1 Introduction

The General Agreement on Tariffs and Trade (GATT), now subsumed by the World Trade Organization (WTO), has governed global trade liberalization since 1948. At the heart of the WTO system is the Most Favored Nation (MFN) principle, contained in Article I of the GATT. This non-discrimination principle requires countries within the WTO to impose the same tariff—the MFN *applied tariff*—on other WTO countries. At the same time however, directly conflicting with its non-discrimination stance, GATT allows discriminatory liberalization through Preferential Trade Agreements (PTAs) under Article XXIV. Specifically, this article allows countries who are members of PTAs to eliminate tariffs between themselves, which they don't have to extend to non-member countries outside of the PTAs, provided they do not increase their tariff barriers on these non-members. Although relatively rare before the Uruguay Round, PTAs have proliferated thereafter. The increasing prevalence of PTAs has generated a large literature that focuses on understanding how the inherently discriminatory nature of PTAs impacts the degree of global trade liberalization that would otherwise arise in terms of non-discriminatory MFN applied tariffs set by countries either individually or through multilateral negotiations.

While PTAs embody discrimination against non-member countries, MFN requires countries not participating in PTAs themselves to treat PTA participants in a non-discriminatory fashion. This awkward asymmetry raises two substantive questions. One, is there a case for allowing PTA non-members to deny MFN treatment to PTA members? In other words, should non-members be permitted to engage in tariff discrimination when they find themselves facing such discrimination at the hands of PTA members? Two, does the answer to this question depend upon whether non-members have *voluntarily* chosen to stay out of a PTA (knowing full-well that their non-participation will result in them facing discriminatory treatment) or have been *deliberately excluded* by PTA members? Our model allows us to directly address these novel questions that have been overlooked in the vast literature on PTAs.

Our formal approach follows Saggi and Yildiz (2010) who develop an equilibrium theory of PTAs in a modified version of the three-country competing exporters framework of Bagwell and Staiger (1999a).[1] Assuming FTA members impose zero tariffs on one another, they compare the relative merits of bilateralism and multilateralism as alternative routes to global trade liberalization. In the present paper, like Saggi and Yildiz (2010), we begin with a *WTO scenario* under which the non-member is obligated to follow the MFN principle of non-discrimination when setting its external tariffs. We next compare this WTO scenario with an alternative setting, called *tariff discrimination scenario.* Under this setting, the

[1] Saggi et. al (2013) build on Saggi and Yildiz (2010) by considering trade agreements that take the form of customs unions as opposed to FTAs.

non-member country does not have to abide by MFN and is free to impose its optimal discriminatory tariffs on PTA members. Intuitively, a comparison of the two scenarios helps determine whether there is a sound rationale for requiring a country to practice non-discrimination when it itself faces discrimination at the hands of PTA members.

From the existing literature we know that optimal MFN tariffs generally impose fewer distortions relative to optimally chosen discriminatory tariffs.[2] In general, since discriminatory tariffs are biased against the efficient exporters, they cause socially harmful trade diversion. A comparison of the WTO-consistent scenario with the tariff discrimination scenario brings to light a hitherto ignored benefit of MFN: by making tariff discrimination infeasible, MFN reduces the potency of a country's optimal tariffs and therefore its incentive for unilaterally opting out of trade liberalization with other countries. Thus, by increasing the likelihood of each country voluntarily choosing to enter into international trade agreements, the MFN principle can act as a catalyst for further trade liberalization. However, we also show that this pro-liberalization effect of MFN is weaker when one country is *deliberately excluded* by the other two (who prefer a bilateral trade agreement with each other to a multilateral one). In other words, we show that the welfare case for requiring a country to follow MFN as a non-member trading with countries that are in a bilateral PTA with each other is stronger if it has voluntarily chosen to not enter into trade agreements with its trading partners relative to a scenario where it has been excluded from their bilateral PTA against its wishes.

In our model, a pair of countries have an incentive to deliberately exclude the third country only when they can coordinate their external tariffs (which is in the case of CUs). The practical implication from this result is that the case for requiring MFN on the part of the CU non-member is relatively weaker than the case for the FTA non-member. This is because the latter is voluntarily choosing to stay out of FTAs in order to benefit from the reductions in the external tariffs of FTA members while retaining the freedom to utilize its optimal discriminatory tariffs.

We also examine the implications of allowing PTA non-members to deny MFN treatment to PTA members in a world where free trade is infeasible because of the underlying economic environment. Under such a case, we find that MFN adoption is world welfare improving regardless of the nature of PTAs. This result extends support for the idea that MFN requirement is not necessarily consistent with the prospect of global free trade while it avoids the socially harmful trade diversion in a tariff ridden world.

Since Bhagwati (1991), a rich literature has emerged to address the question whether PTAs serve as *building* or *stumbling* blocs for multilateral trade liberalization. Early the-

[2]See Choi (1995), Bagwell and Staiger (1999b), Horn and Mavroidis (2001), McCalman (2002), Saggi (2004), and Bagwell and Staiger (2010) for anlyses of the various legal and economic aspects of MFN.

oretical research in this area generally took PTAs to be exogenously given and focused on how PTA membership affects the incentives that countries have for participating in multilateral trade liberalization (see, for example, Krishna, 1998; Ornelas, 2005a, 2005b). The next wave of studies, such as Goyal and Joshi (2006), Aghion et al. (2007), Furusawa and Konishi (2007), and Seidman (2009) consider endogenous PTAs but ignore the possibility of MFN based trade liberalization. A series of papers published in the last decade or so – such as Saggi and Yildiz (2010), Saggi et. al (2013), Missios et al. (2016) and Stoyanov and Yildiz (2015) – have argued that PTAs ought to be seen as building blocks only if the freedom to pursue PTAs (granted to WTO members by GATT Article XXIV) is necessary for achieving global free trade. Additionally, Saggi, Wong, and Yildiz (2019) show that the free internal trade requirement of Article XXIV makes it harder to achieve global free trade, i.e., it reduces the likelihood that PTAs act as building blocs. An attractive feature of this recent line of research is that it treats both preferential and multilateral liberalization as being endogenous. The present paper follows this approach and furthers the literature on the building versus stumbling bloc question by showing that whether or not requiring MFN on the part of the non-member country is conducive for the cause of global free trade depends upon the *nature* of the PTA in question: MFN facilitates free trade when PTAs take the form of FTAs whereas it hinders it if they take the form of CUs.

2 Tariffs and Trade

Our underlying trade model is an appropriately adapted version of the partial equilibrium 'competing exporters' framework developed by Bagwell and Staiger (1999a) to analyze the effects of PTAs. There are three asymmetrically endowed countries: i, j, and k and three (non-numeraire) goods: I, J, and K.[3] Each country's market is served by two competing exporters and I denotes the good that corresponds to the upper case value of i. Country i is endowed with zero units of good I and e_i units of the other two goods.

The demand for good z in country i is given by

$$d(p_i^z) = \alpha - p_i^z \text{ where } z = I, J, \text{ or } K \tag{1}$$

As is well known, the above demand functions can be derived from a utility function of the form $U(c^z) = u(c^z) + w$ where c^z denotes consumption of good z; w denotes the numeraire good; and $u(c^z)$ is quadratic and additively separable in each of the three goods. Country i must import good I in order to consume it and it can import it from either trading partner.

[3] All countries have large enough endowments of the freely traded numeraire good that they consume in positive quantities.

Let t_{ij} be the tariff imposed by country i on its imports of good I from country j. Ruling out prohibitive tariffs yields the following no-arbitrage conditions:

$$p_i^I = p_j^I + t_{ij} \ = p_k^I + t_{ik} \tag{2}$$

Let m_i^I be country i's imports of good I. Since country i has no endowment of good I, we have

$$m_i^I = d(p_i^I) = \alpha - p_i^I \tag{3}$$

Each country's exports of a good must equal its endowment of that good minus its local consumption:

$$x_j^I = e_j - [\alpha - p_j^I] \tag{4}$$

Market clearing for good I requires that country i's imports equal the total exports of the other two countries:

$$m_i^I = \sum_{j \neq i} x_j^I \tag{5}$$

Equations (2) through (5) imply that the equilibrium prices of good I in country i and country j equal:

$$p_i^I = \frac{1}{3}\left(3\alpha - \sum_{j \neq i} e_j + \sum_{j \neq i} t_{ij}\right) \tag{6}$$

$$p_j^I = \frac{1}{3}\left(3\alpha - \sum_{j \neq i} e_j - 2t_{ij} + t_{ik}\right) \tag{7}$$

A country's terms of trade motive for import tariffs is evident from equations (6) and (7): only a third of a given increase in either of its tariffs passes through to a higher local price p_i^I while the remaining two third passes through to a lower exporter price p_j^I which represents a terms of trade improvement for country i vis-à-vis country j.

From a welfare perspective, given the partial equilibrium nature of the model, it suffices to consider only protected goods. A country's welfare is defined as the sum of consumer surplus, producer surplus, and tariff revenue over all such goods:

$$w_i = \sum_z CS_i^z + \sum_z PS_i^z + TR_i \tag{8}$$

Using above equations, one can easily obtain welfare of country i as a function of endowment levels and tariffs. Let aggregate world welfare be defined as the sum of each country's welfare:

$$ww = \sum_i w_i. \tag{9}$$

Before proceeding further, we note that in order to guarantee non-negative exports and positive tariffs under all trade policy regimes, we impose the following parameter restriction throughout the paper: $\max\{e_i, e_j, e_k\} \leq \frac{5}{4}\min\{e_i, e_j, e_k\}$.[4]

We are now ready to report the key properties of the different types of optimal tariffs that arise under the various trade policy regimes that can arise in our model.

Suppose countries do not enter into any type of trade agreement with each other and let t_i^M denote country i's optimal MFN tariff:

$$t_i^M \equiv Arg\max w_i(\Phi) = \frac{e_j + e_k}{8}$$

Upon forming an FTA, member countries remove their internal tariffs on each other and impose an individually optimal external tariff on the non-member. Under a single FTA, say between i and j, the optimal external tariff of an FTA member is

$$t_{ik}(ij) = \frac{5\,e_k - 4e_j}{11} \tag{10}$$

Comparing $t_{ik}(ij)$ and t_i^M reveals that $t_i^M > t_{ik}(ij)$. That is, FTA members practice "tariff complementarity": FTA formation induces them to lower their tariff on the FTA outsider.[5] Conversely, market separability implies that FTA formation leaves the FTA non-member's optimal MFN tariff unchanged: $t_{ki}(ij) = t_{kj}(ij) = t_k^M$.

Next, consider a bilateral CU formation. Like FTA members, CU members remove internal tariffs on each other but, unlike FTA members, CU insiders coordinate external tariffs. Maximizing their joint welfare, their optimal external tariff is

$$t_{ik}(ij^u) \equiv \underset{t_{ik}(ij^u)}{\arg\max}\; w_i(ij^u) + w_j(ij^u) \text{ subject to } t_{jk}(ij^u) = t_{ik}(ij^u) \tag{11}$$

$$= \frac{2\,e_k - e_j}{5}. \tag{12}$$

While CU members also practice tariff complementarity, i.e. $t_i^M > t_{ik}(ij^u)$, their degree of

[4] Calculations supporting this restriction and all of the results reported in the paper are contained in the appendix.

[5] For tariff complementarity discussions, see Bagwell and Staiger (1997, 1999), Bond (2004), Saggi and Yildiz (2009) and Estevedeordal (2008). While only terms-of-trade considerations influence trade policy considerations here, this result is robust to including other trade policy motives like the presence of a production relocation externality (Suwanprasert (2018) and Ossa (2011))

tariff complementarity is less than FTA insiders: $t_i^M - t_{ik}(ij^u) < t_i^M - t_{ik}(ij)$.[6] Intuitively, when setting external tariffs individually, each FTA member ignores the negative externality imposed on the export surplus of its FTA partner by lowering its tariff on the FTA non-member. By coordinating their external tariffs, CU members internalize this negative externality, i.e. $t_{ik}(ij^u) > t_{ik}(ij)$, and thereby benefit from tariff coordination.

Here, one should note that, in the absence of an MFN clause, a non-member country is free to tariff discriminate across its trading partners. For a non-member country (say country i) under an FTA $\langle jk \rangle$, (t_{ij}^*, t_{ik}^*) solve $\arg\max w_i(jk)$:

$$t_{ij}^*(jk) = \frac{3e_j - e_k}{8} \quad \text{and} \quad t_{ik}^*(jk) = \frac{3e_k - e_j}{8}$$

Note that, when free to tariff discriminate, each country imposes a higher tariff on the larger exporter:

$$t_{ij}^*(jk) - t_{ik}^*(jk) = \frac{(e_j - e_k)}{2} > 0 \text{ iff } e_j > e_k$$

By increasing its volume of imports, a country's optimal tariff increases with the exporters' endowments. Relative to the case of MFN, countries impose higher discriminatory tariff on the imports from the country with larger endowment:

$$t_i^M - t_{ij}^*(\Phi) = \frac{2(e_k - e_j)}{8} < 0 \text{ iff } e_j > e_k$$

and that each country's optimal MFN tariff is bound by its discriminatory tariffs:

$$t_{ij}^*(jk) \leq t_i^M \leq t_{ik}^*(jk) \text{ where } e_j \leq e_k \tag{13}$$

We summarize the key messages of the above analysis in the following lemma is well-established in the literature (see for example, Saggi (2004)):

Lemma 1 *When tariff discrimination is allowed, the non-member country under a bilateral FTA imposes a higher tariff on the country from which it sources a larger volume of import: $t_{ik}^*(jk) \geq t_{ij}^*(jk)$ iff $e_j \leq e_k$ while the non-member country's optimal MFN tariff is bound by its optimal discriminatory tariffs: $t_{ij}^*(jk) \leq t_i^M(jk) \leq t_{ik}^*(jk)$ when $e_j \leq e_k$.*

One should note here that, the non-member country under a bilateral PTA always have an incentive to discriminate external tariffs and since the discriminatory tariffs are

[6] In contrast, see Zissimos (2012) for a setting, and the implications thereof, where CU members are indeed bound by the GATT Article XXIV constraint that they do not raise their tariffs on non-members.

biased against the import from the country with higher comparative advantage, it leads to a socially harmful trade diversion. Therefore, in a tariff ridden world under a given trade regime, MFN adoption by each country improves world welfare by eliminating this trade diversion. Later, we will stress on this result in the cases where free trade fails to obtain and bilateral PTAs arise as the equilibrium agreements.

3 Endogenous Free Trade Agreements

The two policy scenarios that we contrast are formalized as follows:

(a) *WTO scenario*: This scenario is captured by a three stage game of trade liberalization under which countries abide by both Article I and Article XXIV of GATT. In the first stage, countries enter into FTAs with one another (the process of FTA formation is described in greater detail below). In the second stage, given the trade policy regime that results from the first stage, countries choose their tariffs. If an FTA is formed, its members eliminate the internal tariffs between themselves while imposing individually optimal external tariffs on the non-member who, in accordance with MFN, imposes non-discriminatory tariffs on the two member countries. At the third stage of the game, given trade agreements and tariffs, international trade and consumption take place.

(b) *Tariff discrimination scenario*: This scenario differs from the WTO-consistent benchmark in one way: at the second stage of the game, the non-member country is free to impose discriminatory tariffs on FTA members as opposed to having to treat them in an MFN manner. Thus, all countries engage in some type of tariff discrimination: FTA members discriminate against the non-member by imposing higher tariffs on it than they do on each other while the non-member discriminates between them by imposing a higher tariff on the country from whom it imports more (see Lemma 1).

We now describe the process of FTA formation that occurs during the first stage of the game and is common to all three scenarios.

The process of FTA formation: At the first stage of the game, each country announces whether or not it wants to sign an FTA with each of the other two countries. Denote country i's announcement by σ_i and its strategy set by S_i where

$$S_i = \{\{\phi, \phi\}, \{j, \phi\}, \{\phi, k\}, \{j, k\}\} \tag{14}$$

In S_i, $\{\phi, \phi\}$ denotes an announcement in favor of no FTAs, $\{j, \phi\}$ an announcement in favor of an FTA with only country j; $\{\phi, k\}$ in favor of an FTA with only country k; and $\{j, k\}$ in favor of FTAs with both of them. Since a trade agreement requires consent from

both sides, we posit the following mapping between various announcements profiles and the types of trade agreements that countries can form:

(i) No two announcements match or the only matching announcements are $\{\phi, \phi\}$. All of these announcement profiles yield no agreement $\langle\Phi\rangle$. Under both scenarios, all countries impose their optimal MFN tariffs on one another.

(ii) Two countries announce each others' name and there is no other matching announcement: i.e., $j \in \sigma_i$ and $i \in \sigma_j$ while $i \notin \sigma_k$ and/or $k \notin \sigma_i$ and $j \notin \sigma_k$ and/or $k \notin \sigma_j$. All of these announcements yield an FTA between countries i and j denoted by $\langle ij\rangle$ under which members impose individually optimal external tariffs on the non-member who imposes optimal MFN tariff on members under the WTO scenario while imposing optimal discriminatory Nash tariffs under the tariff discrimination scenario.

(iii) Country i announces in favor of signing an FTA with countries j and k while countries j and/or k announce only in favor of signing an FTA with country i: i.e. $\sigma_i = \{j, k\}$; $i \in \sigma_j$; and $i \in \sigma_k$ while $k \notin \sigma_j$ and/or $j \notin \sigma_k$. This set of announcements yields a pair of independent FTAs (i.e. a hub and spoke trading regime) with i as the common member denoted by $\langle ij, ik\rangle$ (or simply $\langle ih\rangle$).

(iv) All countries announce each others' names, i.e., the announcement profile is $\Omega^F \equiv \{\sigma_i = \{j, k\},\ \sigma_j = \{i, k\},\ \sigma_k = \{i, j\}\}$. This announcement profile yields global free trade $\langle F\rangle$.

Note that since an FTA between two countries can arise only if it is mutually acceptable to both sides, multiple announcement profiles can map into the same agreement. For example, the FTA $\langle ij\rangle$ obtains when (i) countries i and j call only each other, *regardless* of the nature of country k's announcement: if $\sigma_i = \{j, \phi\}$ and $\sigma_j = \{i, \phi\}$, then $\langle ij\rangle$ is the result of all four possible announcements on the part of country k, i.e., for $\sigma_k = \{\phi, \phi\}$, $\{i, \phi\}$, $\{\phi, j\}$ and $\{i, j\}$ so that country k's announcement has no bearing upon the outcome when neither of the other two countries' announce its name; (ii) countries i and j announce each other's name and either one or both of them also announce country k but country k does not reciprocate: i.e. all of the following types of announcements map into the FTA $\langle ij\rangle$: (a) $\sigma_i = \{j, k\}$ and $\sigma_j = \{i, \phi\}$ but $i \notin \sigma_k$, or (b) $\sigma_i = \{j, \phi\}$ and $\sigma_j = \{i, k\}$ but $j \notin \sigma_k$, or (c) $\sigma_i = \{j, k\}$ and $\sigma_j = \{i, k\}$ but $\sigma_k = \{\phi, \phi\}$. As in Saggi and Yildiz (2010), Saggi et al. (2013) and Missios et al. (2016), we can easily rule out all the non-parsimonious announcements as candidates for Nash equilibria and thus we will focus on the parsimonious announcements only.

When analyzing the above games, we refine the set of Nash equilibria by isolating

those Nash equilibria that are *coalition proof*. Bernheim et al. (1987) state that "in an important class of "noncooperative" environments, it is natural to assume that players can freely discuss their strategies, but cannot make binding commitments. In such cases, any meaningful agreement between the players must be self-enforcing. Although the Nash best-response property is a necessary condition for self-enforceability, it is not sufficient - it is in general possible for coalitions to arrange plausible, mutually beneficial deviations from Nash agreements." Therefore, a coalition proof Nash equilibrium (CPNE) is a Nash equilibrium that is immune to all *self-enforcing* coalitional deviations.

4 Equilibrium Free Trade Agreements

In order to simplify exposition, we make the following assumption:

Assumption 1 *There are three asymmetric countries. Country l is the largest importer relative to countries m and s while medium country m is a larger importer than small country s: $e_s = \theta e \geq e_m = \frac{1}{2}\theta e + \frac{1}{2}e \geq e_l = e$ where $1 \leq \theta \leq 5/4$.*

It is worth pointing out here that, in our model, all countries have the ability to manipulate their terms of trade via import tariffs. Country s has a weaker ability to manipulate its terms of trade but its not a 'small' country in the traditional sense of the term wherein it would be a price-taker on world markets.

We proceed as follows. First, we study FTA formation in our WTO-consistent benchmark scenario and show that, while two countries may have an incentive to form a bilateral trade agreement aimed at excluding the third country, this exclusion incentive goes unexercised since it is not self-enforcing. Instead, it is the strength of the free-riding incentive of the non-member country that proves pivotal in determining whether or not global free trade obtains in equilibrium. Then, we argue that the ability of the non-member country under a bilateral FTA to discriminate external tariffs makes it harder to achieve global free trade due to larger free riding incentive relative to the WTO consistent world where Article I binds. Later, with our CU game we show that both the equilibrium and welfare implications of the MFN clause depend on whether the non-member voluntarily stays out of the PTA between the other two countries or has been deliberately excluded by them.

4.1 WTO-consistent agreements

In this section, we derive equilibrium trade agreements under our benchmark scenario where countries follow both Articles I and XXIV of GATT – i.e. the non-member country follows

MFN and FTA members engage in free internal trade and do not raise their external tariffs on non-members. Let country i's welfare as a function of the underlying trade policy regime r be denoted by $w_i(r)$, where $r = \langle\Phi\rangle$,$\langle ij\rangle$, $\langle ih\rangle$, or $\langle F\rangle$ and it is understood that all countries impose optimal tariffs consistent with regime r. Let $\Delta w_i(r-v)$ denote the difference between country i's welfare under trade agreements r and v: $\Delta w_i(r-v) \equiv w_i(r) - w_i(v)$, where $r, v = \langle\Phi\rangle$,$\langle ij\rangle$, $\langle ih\rangle$, or $\langle F\rangle$. Furthermore, let $\theta_i(r-v)$ denote the critical threshold of asymmetry at which country i is indifferent between regimes r and v.

We first state the following lemma that explains how differences in endowment across countries lead them to have asymmetric preferences over various trade regimes:

Lemma 2 *In the WTO-consistent approach to the formation of trade agreements, the following holds:*

(i) (Attractiveness of bilateral FTA formation) *Under no agreement, countries always have an incentive to form a bilateral FTA and each country prefers to form a bilateral FTA with the larger importer relative to the smaller one:* $\Delta w_l(ml-sl) > 0$, $\Delta w_m(ml-sm) > 0$, *and* $\Delta w_s(sl-sm) > 0$ *for all* θ.

(ii) (Free riding incentive) *While small and medium countries do not have an incentive to free ride and stay an outsider under a bilateral FTA relative to free trade, the large importer does have such incentive when the degree of endowment asymmetry is sufficiently large:* $\Delta w_i(F-jk) > 0$ *for all* θ *where* $i = s, m$ *and* $i \neq j, k$ *while* $\Delta w_l(F-sm) < 0$ *when* $\theta > \theta_l(F-sm)$.

(iii) (Exclusion incentives) *While small importing country does not participate in any joint deviation from free trade to exclude the third country via a bilateral FTA, medium and large importers do have such incentives when the degree of endowment asymmetry is sufficiently large:* $\Delta w_s(F-sj) > 0$ *for all* θ *where* $j = m, l$ *while* $\Delta w_i(F-ij) < 0$ *when* $\theta > \theta_i(F-ij)$ *where* $i, j = m, l$.

(iv) (Attractiveness of being hub) *All countries prefer being the hub under a hub and spoke regime relative to all other trade policy regimes:* $\Delta w_i(ih-\Phi) > 0$; $\Delta w_i(ih-F) > 0$ *and* $\Delta w_i(ih-ij) > 0$ *for all* $i = s, m, l$.

(v) (Attractiveness of being spoke) *While the large importer always prefers being a non-member under a bilateral FTA to being a spoke under a hub and spoke regime, the medium and smaller importers do so only when the degree of endowment asymmetry is sufficiently small:* $\Delta w_l(ih-sm) < 0$ *for all* θ *and* $i = s, m$ *and* $\Delta w_i(jh-jk) < 0$ *when* $\theta < \theta_i(jh-jk)$ *where* $i = s, m$ *and* $i \neq j, k$.

Part (i) of Lemma 2 governs the attractiveness of bilateral FTA formation when there exists no FTA. Intuitively, the exchange of preferential access makes bilateral FTA formation attractive. The latter part follows from two reinforcing effects. The larger a country's trading partner's import volume, the larger the increase in export surplus it enjoys from the elimination of its partner's optimal tariff and the smaller the loss it suffers from its own trade liberalization since its tariff reduction applies to a smaller volume of imports. Thus, a country prefers to form a bilateral FTA with the larger importer amongst its two trading partners.

The second part of Lemma 2 argues that small and medium exporters have no incentive to unilaterally deviate from free trade to become an outsider under a bilateral FTA. However, when the large importer is sufficiently large, it has an incentive to free ride on the trade liberalization by the other two countries. Intuitively, despite the discrimination faced as an FTA outsider, it benefits from tariff complementarity which lowers the external tariffs faced when exporting to the FTA insiders and retains its ability to impose optimal tariffs.

The second and third parts of the Lemma together inform us that the small importing country has strong preference for freer trade and thus has no incentive to unilaterally or jointly deviate from free trade. This is because its volume of export is large while its volume of import is small relative to other countries. However, medium and large importers have incentives to jointly exclude the small importer from their free trade network when the asymmetry is sufficiently large.

The fourth part of the Lemma says that being a hub country is preferable for all countries relative to other trade policy regimes irrespective of their relative sizes. Intuitively, the hub country enjoy sole preferential access in the spoke countries while offering free trade in the domestic market. Note in particular that, relative to free trade, the hub country enjoys privileged access in both spoke countries while its domestic surplus is no different. Moreover, this privileged access in export markets is so desirable that a hub country has no incentive to unilaterally revoke either or both of its FTAs.

Finally, under a bilateral FTA, the large importer has no incentive to give up its ability to impose optimal tariffs to gain a free access in an export market in which the competing exporter already has a free access. This implies that hub and spoke regimes in which small and medium countries are hub are not even Nash equilibrium.

An important message delivered by the above lemma is that the small importing country's preference does not matter for the equilibrium condition of a trade agreement while

the larger importing country's choice is critical. We should also note from the first part of the Lemma 2 that countries always have an incentive to form a bilateral FTA. While members of an FTA discriminate against the non-member country, we know from the above tariff analysis that tariff complementarity partially benefits the non-member who also retains its ability to impose optimal external MFN tariffs. This raises the possibility that, starting from no agreement $\langle\Phi\rangle$, the formation of an FTA makes all countries better off (i.e. is Pareto improving relative to $\langle\Phi\rangle$). Indeed, we can show that the medium and large importing countries always benefits from the formation of a bilateral FTA in which they are not member of regardless of the degree of asymmetry:

$$\Delta w_i(jk - \Phi) > 0 \textit{ for all } \theta,\ i = m, l \textit{ and } i \neq j, k \tag{15}$$

Given its strong preferences for better export access, the small importing country benefits from the formation of the bilateral FTA between larger importers only when the degree of endowment asymmetry is sufficiently small:

$$\Delta w_s(ml - \Phi) \geq 0 \textit{ when } \theta \leq \theta_s(ml - \Phi)$$

Therefore, we find the following:

Proposition 1 *Relative to no agreement $\langle\Phi\rangle$, bilateral FTAs $\langle sm\rangle$ and $\langle sl\rangle$ are Pareto-improving for all θ while the the FTA $\langle ml\rangle$ is Pareto-improving iff $\theta \leq \theta_s(ml - \Phi)$.*

We are now ready to determine the CPNE of the FTA formation game under the WTO consistent scenario. We proceed by considering each of the announcement profiles that yield the various trade policy regimes in turn. First, consider the announcement profile leading to global free trade $\langle F\rangle$. First note from parts (ii) and (iii) of the Lemma 2 that small importer (i.e. country s) has no incentive to participate in any deviation (unilateral or coalitional). Thus, if there exists a coalitional deviation, it must involve countries m and l. It is immediate from part (iii) of the above lemma that, taking country s' announcement fixed at $\{m, l\}$, countries m and l have an incentive to jointly deviate from their respective announcements $\{s, l\}$ and $\{s, m\}$ to $\{\phi, l\}$ and $\{\phi, m\}$ in order to exclude country s from a free trade network when country s is sufficiently small:

$$\Delta w_l(F - ml) \quad < \quad 0 \textit{ when } \theta > \theta_l(F - ml) \tag{16}$$

$$\Delta w_m(F - ml) \quad < \quad 0 \text{ when } \theta > \theta_m(F - ml) \tag{17}$$

Since $\theta_m(F-ml) > \theta_l(F-ml)$, $\theta_m(F-ml)$ is the binding cutoff for this joint deviation. The above result establishes the existence of an *exclusion incentive*: when the endowment asymmetry is sufficiently pronounced (i.e. $\theta > \theta_m(F-ml)$) the two larger importers prefer a bilateral FTA between themselves relative to global free trade. The key question is whether the joint exclusion incentive of the two larger importers is self-enforcing or not. The answer to this question turns out to be negative. To see why, suppose each country announces in favor of an FTA with both its trading partners. Starting with these announcements the two larger importers have an incentive to exclude the smaller country by jointly altering their announcements such that the announcement profile changes from Ω^F (which yields free trade) to $\Omega_1^{ml} = \{\sigma_l = \{\phi, m\}, \sigma_m = \{\phi, l\}, \sigma_s = \{m, l\}\}$ thereby altering the associated trade regime from free trade to the bilateral FTA $\langle ml \rangle$. However, from part (iv) of Lemma 2 we know that each country's most preferred trading arrangement is a hub and spoke regime with itself as the hub. It follows then that, holding constant the announcement of the excluded country at $\sigma_s = \{m, l\}$, each member of the initially deviating coalition (m or l) has an incentive to alter its announcement so as to include country s. For example, country l has an incentive to alter its announcement from $\sigma_l = \{\phi, m\}$ to $\sigma_l = \{s, m\}$ which alters the trade regime from $\langle ml \rangle$ to $\langle lh \rangle$. Since the welfare of a hub is higher than that of a member country in a single FTA – see part (iv) of Lemma 2 – the original coalitional deviation of countries m and l from Ω^F to Ω_1^{ml} is *not* self-enforcing. Thus, in a nutshell, the lure of a hub and spoke trading arrangement makes any joint deviation from Ω^F to an announcement profile that supports a bilateral FTA not-self enforcing.

Since all countries are better off under free trade relative to $\langle \Phi \rangle$, joint announcement deviations that convert the trade regime from $\langle F \rangle$ to $\langle \Phi \rangle$ never arise. Based on the above discussion and the lemma, the only possible type of self-enforcing deviations from Ω^F that we need to consider are unilateral deviations from Ω^F and they are self-enforcing by definition. First, we establish that small and medium countries have no incentives to unilaterally deviate from Ω^F to announcements leading to hub and spoke regimes in which they are spokes. How about the large importing country? We find that, when the degree of asymmetry is sufficiently large, it has an incentive to unilaterally deviate from its announcements $\{s, m\}$ to $\{s, \phi\}$ or $\{\phi, m\}$ that leads to a hub and spoke regime under which countries s or m is a hub and it itself is a spoke:

$$\begin{aligned} \Delta w_l(F-sh) &< 0 \text{ when } \theta > \theta_l(F-sh) \\ \Delta w_l(F-mh) &< 0 \text{ when } \theta > \theta_l(F-mh) \end{aligned}$$

where $\theta_l(F - sh) > \theta_l(F - mh)$. Then, it is immediate from the second part of the above lemma that only one unilateral deviation incentive remains to be examined: unilateral deviation of country l unilaterally from $\{s, m\}$ to $\{\phi, \phi\}$ converting free trade to $\langle sm \rangle$ and it happens when country l is sufficiently large importer relative to other countries:

$$\Delta w_l(F - sm) < 0 \text{ when } \theta > \theta_l(F - sm) \tag{18}$$

We find that $\theta_l(F - sm) < \theta_l(F - mh)$ and thus the announcement profile leading to $\langle F \rangle$ is CPNE whenever $\theta \leq \theta_l(F - sm)$.

What if $\langle F \rangle$ is not a CPNE, as is the case when $\theta > \theta_l(F - sm)$? We can quickly rule out the various announcement profiles leading to the hub and spoke regimes as candidates for CPNE. To see why, recall from part (v) of Lemma 2 that the large importing country under $\langle sh \rangle$ and $\langle mh \rangle$ always has an incentive to unilaterally deviate from its respective announcements $\{s, \phi\}$ and $\{\phi, m\}$ to $\{\phi, \phi\}$ and $\{\phi, \phi\}$, leading to a deviation from $\langle sh \rangle$ to $\langle sm \rangle$ and from $\langle mh \rangle$ to $\langle sm \rangle$. Since these unilateral deviations are self-enforcing, any announcement profile leading to $\langle sh \rangle$ and $\langle mh \rangle$ is not even a Nash equilibrium (thus cannot be a CPNE). Consider now the announcement profile that leads to $\langle lh \rangle$. Smaller importing spoke countries always have an incentive to jointly deviate from their announcement to form the final FTA leading to global free trade and it is immediate from the above lemma that neither of these countries have incentive to unilaterally deviate further. Thus the initial deviation is self enforcing and the announcement profile that leads to $\langle lh \rangle$ is not a CPNE.

Next, we consider the announcement profile that leads to no agreement $\langle \Phi \rangle$. The first part of our Lemma 2 informs us that two countries always have an incentive to deviate from their announcements to the ones that lead to a bilateral FTA and this joint announcement deviation is self-enforcing. As a result, the announcement profile that yields $\langle \Phi \rangle$ cannot be a CPNE.

The only remaining candidates for CPNE are the announcement profiles that lead to bilateral FTAs. We start with those profiles that yield an FTA between the small and medium importers, say $\langle sm \rangle$). Note from part (ii) and part (v) of Lemma 2 that country l has no incentive to unilaterally deviate from its announcement $\{\phi, \phi\}$ to either include the small or medium country or both when $\theta > \theta_l(F - sm)$. Under such a case, the small and medium countries also have no incentive to break up their mutual FTA. As a result the announcement profile that yields $\langle sm \rangle$ is a CPNE when $\theta > \theta_l(F - sm)$ holds.

Next, we move to announcement profiles that yield $\langle sl \rangle$. We find that neither country has an incentive to unilaterally break up their agreements (see Proposition 1). Second, we

know from part (i) of Lemma 2 that country s has no incentive to engage in the coalitional announcement deviation with country m that converts $\langle sl \rangle$ to $\langle sm \rangle$. Third, coalitional announcement deviation of countries m and l converting $\langle sl \rangle$ to $\langle ml \rangle$ is *not* self-enforcing since the common member country (i.e. country l) has an incentive to further deviate to become the hub country, taking the announcement of its complement as fixed. Fourth, note from the above discussion that the coalitional announcement deviation that replaces $\langle sl \rangle$ by $\langle F \rangle$ is self-enforcing only when $\theta \leq \theta_l(F - sm)$. Finally, we know from part (iv) of Lemma 2 that small and large importing countries always have an incentive to become hub and it is immediate from part (v) of Lemma 2 that country m has an incentive to engage in any coalitional announcement deviations that replace $\langle sl \rangle$ by $\langle sh \rangle$ or $\langle sl \rangle$ by $\langle lh \rangle$ when the degree of endowment asymmetry is sufficiently large: $\theta > \theta_m(sh - sl) > \theta_m(lh - sl)$. These deviations are self-enforcing since neither country has an incentive to unilaterally deviate further. As a result, the announcement profile leading to $\langle sl \rangle$ is a CPNE whenever $\theta_l(F - sm) \leq \theta \leq \theta_m(lh - sl)$.

Finally, we consider the bilateral FTA between the two larger importing countries, i.e., $\langle ml \rangle$. First, as before, the coalitional announcement deviation from $\langle ml \rangle$ to $\langle F \rangle$ occurs $\theta \leq \theta_l(F - ml)$ and it is self-enforcing when $\theta \leq \theta_l(F - sm)$. Second, we can show that when $\theta > \theta_s(lh - ml)$ country s and country l have an incentive to jointly deviate from their respective announcements $\{\phi, \phi\}$ and $\{\phi, m\}$ to $\{l, \phi\}$ and $\{s, m\}$, leading to a deviation from $\langle ml \rangle$ to $\langle lh \rangle$ and this deviation is self-enforcing. Since $\theta_s(lh - ml) < \theta_l(F - sm)$, these self-enforcing announcement deviations cover the entire parameter space and thus the announcement profile supporting $\langle ml \rangle$ is never a CPNE.

We summarize the main findings of the above analysis below:

Proposition 2 *The equilibria of the WTO-consistent game of trade liberalization where FTA members have to practice free internal trade and the non-member has to abide by MFN are as follows:*

(i) Free trade $\langle F \rangle$ is the unique CPNE when $\theta \leq \theta_l(F - sm)$.[7]

(ii) Bilateral FTAs $\langle sl \rangle$ and $\langle sm \rangle$ are CPNE when $\theta_l(F - sm) \leq \theta \leq \theta_m(lh - sl)$.

(iii) Bilateral FTA $\langle sm \rangle$ is the unique CPNE when $\theta > \theta_m(lh - sl)$.

[7] We should note here that, technically speaking, the equilibrium is the announcement profile Ω^F that yields free trade as the agreement. In what follows, for expositional ease, we state our results directly in terms of various trade agreements that emerge as equilibrium outcomes as opposed to the announcement profiles that support them.

Insert Figure 1

The above proposition relates the degree of underlying asymmetry to the nature of equilibrium agreements. Part (i) simply says that if the degree of endowment asymmetry is sufficiently small, free trade is the equilibrium outcome. One important insight that emerges from the above discussion is that exclusion incentives go unexercised due to the lure of attractive hub and spoke trading arrangements and free riding incentive of the large importing country is pivotal for the stability of global free trade. Part (ii) says that if the degree of endowment asymmetry is sufficiently large, two asymmetric FTAs ($\langle sl \rangle$ or $\langle sm \rangle$) are the equilibrium outcomes – in both situations, one of the larger importing countries prefers being a non-member to participating in any bilateral or multilateral agreements. Under such a case, our theory offers no guidance regarding which of the trade regimes should be expected to arise in equilibrium. Note also from the above discussion that the bilateral FTA between the two larger countries $\langle ml \rangle$ fails to arise in equilibrium. Finally, when the degree of asymmetry is sufficiently large, part (iii) of Proposition 2 informs us that only the FTA between small and medium importing countries arises as a CPNE.

4.2 Equilibrium agreements under tariff discrimination

While Article XXIV of GATT sanctions tariff discrimination in the form of FTAs, the MFN principle requires WTO members to treat all other members (with whom they do not have FTAs) in a non-discriminatory fashion. A natural question is whether there is a case for imposing the MFN constraint on the tariff-setting behavior of countries trading with FTAs whose members actively discriminate against them. This concern would appear to be even more acute for the case of a country that finds itself deliberately excluded from an FTA. To address these issues, we now analyze our *tariff discrimination* scenario under which not only the FTA members discriminate against the non-member but also the non-member country trading with member countries of an FTA is free to set its optimal discriminatory tariffs on them as opposed to having to treat them in a non-discriminatory fashion.

The structure of a country's optimal discriminatory tariffs and their relationship to its optimal MFN tariff is described in Lemma 1. As this lemma indicates, when free to tariff discriminate across its trading partners, a country sets a higher tariff on the country from whom it sources a larger volume of imports (i.e. the one who has a greater comparative advantage).

Let country i's welfare as a function of the underlying trade agreement r optimal

discriminatory tariffs on the part of non-member be denoted by $w_i(\widetilde{r})$. As before, let $\Delta w_i(\widetilde{r} - \widetilde{v}) \equiv w_i(\widetilde{r}) - w_i(\widetilde{v})$. Here, it is important to note that only the external tariffs of the non-member country under a bilateral FTA are different relative to the benchmark WTO consistent scenario. Therefore, the welfare levels change only under those regimes ($\widetilde{\langle ml \rangle}$, $\widetilde{\langle sl \rangle}$, and $\widetilde{\langle sm \rangle}$).

We next examine how Lemma 2 changes when the non-member is free to discriminate the imports from member countries relative to the case where it abides the MFN clause. It is immediate from our Lemma 1 that the non-member country and the member country with a weaker comparative advantage would prefer the case when the non-member country is able to discriminate relative to the case of MFN while the opposite obtains for the member country with a greater comparative advantage:

$$\Delta w_s(\widetilde{sm} - sm) \leq 0; \Delta w_s(\widetilde{sl} - sl) \leq 0 \text{ and } \Delta w_m(\widetilde{ml} - ml) \leq 0 \text{ for all } \theta \quad (19)$$

$$\text{and } \Delta w_i(\widetilde{jk} - jk) \geq 0 \text{ for all } \theta \text{ and } i = s, m, l \textit{ and } i \neq j, k \quad (20)$$

while

$$\Delta w_m(\widetilde{sm} - sm) \geq 0; \Delta w_l(\widetilde{sl} - sl) \geq 0 \text{ and } \Delta w_l(\widetilde{ml} - ml) \geq 0 \text{ for all } \theta \quad (21)$$

We first show that the first part of Lemma 2 continues to hold: countries always have an incentive to form a bilateral FTA and each country prefers to form a bilateral FTA with the larger importer relative to the smaller one. Moreover, medium and large importing countries still always benefit from the formation of a bilateral FTA in which they are not member of regardless of the degree of asymmetry while the small importing country benefits only when the degree of asymmetry is not sufficiently high. Here, one should note that since $\Delta w_s(\widetilde{ml} - ml) \geq 0$ obtains for all θ, the parameter range over which the FTA between medium and large importing countries is Pareto-improving expands. However, one should note that this positive result comes at the expense of lower world welfare under $\left\langle \widetilde{ml} \right\rangle$ relative to $\langle ml \rangle$. Thus, the following is immediate:

Proposition 3 *Relative to no agreement $\langle \Phi \rangle$, bilateral FTAs $\langle \widetilde{sm} \rangle$ and $\left\langle \widetilde{sl} \right\rangle$ are Pareto-improving for all θ while the the FTA $\left\langle \widetilde{ml} \right\rangle$ is Pareto-improving iff $\theta \leq \theta_s(\widetilde{ml} - \Phi)$ where $\theta_s(ml - \Phi) \leq \theta_s(\widetilde{ml} - \Phi)$.*

Next, we examine how the second part of Lemma 2 adapts under the tariff discrimination scenario. We still obtain that small and medium countries do not have an incentive to

free ride and stay an outsider under a bilateral FTA relative to free trade. As we know from our discussion under WTO consistent scenario, the large importing country's choice is pivotal for global free trade to be the CPNE. We know from the inequalities in (20) that the large importing country has a stronger free riding incentive (stronger unilateral incentive to deviate from Ω^F) when it is free to discriminate as a non-member country relative to the MFN case: $\Delta w_l(F - \widetilde{sm}) < 0$ obtains *when* $\theta > \theta_l(F - \widetilde{sm})$ holds and $\theta_l(F - \widetilde{sm}) < \theta_l(F - sm)$.

We also find that, regardless of whether the non-member country imposes MFN or discriminatory tariffs, being hub is very attractive due to sole preferential access in spoke markets. Therefore, even when exclusion incentive arises, it goes unexercised in the equilibrium as it was under the WTO consistent scenario. Finally, since non-member country always benefits from being able to discriminate the external tariffs, countries' incentive to stay as outsider rather than becoming a spoke strengthens under discrimination relative to MFN.

In the light of the above discussion, we first argue that $\langle\Phi\rangle$ is never a CPNE since any two countries have incentives to jointly deviate and form a bilateral FTA and this deviation is self-enforcing. Second, as before under MFN (even stronger under discrimination), we know a large importing spoke country under $\langle sh\rangle$ and $\langle mh\rangle$ always has an incentive to unilaterally deviate from its respective announcements $\{s, \phi\}$ and $\{\phi, m\}$ to $\{\phi, \phi\}$ and $\{\phi, \phi\}$, leading to a deviation from $\langle sh\rangle$ to $\widetilde{\langle sm\rangle}$ and from $\langle mh\rangle$ to $\widetilde{\langle sm\rangle}$. Since unilateral deviations are self-enforcing, the announcement profiles leading to $\langle sh\rangle$ and $\langle mh\rangle$ are never a CPNE. Consider now the announcement profile that leads to $\langle lh\rangle$. As before, spoke countries s and m have an incentive to jointly deviate from their announcements to form the final FTA leading to global free trade and neither country has an incentive to unilaterally deviate further . Thus the initial deviation is self enforcing and the announcement profile that leads to $\langle lh\rangle$ is not a CPNE.

Third, the coalitional announcement deviation from $\left\langle \widetilde{ml} \right\rangle$ to $\langle F\rangle$ happens when $\theta \leq \theta_l(F - \widetilde{ml})$ and it is self-enforcing only when $\theta \leq \theta_l(F - \widetilde{sm})$. Second, we show that, when $\theta > \theta_s(lh - \widetilde{ml})$ holds, country s and the large country l have an incentive to jointly deviate from their respective announcements $\{\phi, \phi\}$ and $\{\phi, l\}$ to $\{\phi, m\}$ and $\{s, m\}$, leading to a deviation from $\left\langle \widetilde{ml} \right\rangle$ to $\langle lh\rangle$ and this deviation is self enforcing. Since $\theta_s(lh - \widetilde{ml}) < \theta_l(F - \widetilde{sm})$, these self-enforcing announcement deviations cover the entire parameter space and thus the announcement profile leading to $\left\langle \widetilde{ml} \right\rangle$ is never a CPNE.

Based on the above discussion, the only possible announcement profiles that can be

CPNE are the ones leading to $\langle F\rangle$, $\widetilde{\langle sl\rangle}$, and $\widetilde{\langle sm\rangle}$. As in the WTO consistent scenario, we can immediately argue that the unilateral deviation of the large importing country from $\{s, m\}$ to $\{\phi, \phi\}$ is critical for the CPNE condition of global free trade:

$$\Delta w_l(F - \widetilde{sm}) < 0 \text{ when } \theta > \theta_l(F - \widetilde{sm})$$

and thus the announcement profile leading to $\langle F\rangle$ is CPNE whenever $\theta \leq \theta_l(F - \widetilde{sm})$. What if $\theta > \theta_l(F - \widetilde{sm})$ holds and global free trade fails to be a CPNE? The only remaining candidates for CPNE are the announcement profiles that lead to bilateral FTAs $\langle \widetilde{sm}\rangle$ and $\left\langle \widetilde{sl}\right\rangle$. We start with the announcement profiles that yields $\langle \widetilde{sm}\rangle$. First, note that country l has no incentive to engage in a deviation from its announcement $\{\phi, \phi\}$ to either include the small or medium countries or both when $\theta > \theta_l(F - \widetilde{sm})$. Moreover, the small and medium importing countries also have no incentive to unilaterally deviate from their announcements leading to $\langle \widetilde{sm}\rangle$ to the one that yields $\langle \Phi\rangle$. Therefore, $\langle \widetilde{sm}\rangle$ is a CPNE when $\theta \geq \theta_l(F - \widetilde{sm})$ holds.

Next we move to profiles that yield an FTA between the small and large importers, $\left\langle \widetilde{sl}\right\rangle$. We know from the above discussion that that neither country has an incentive to unilaterally deviate from its announcement to the one leading to $\langle \Phi\rangle$. We also know that the coalitional announcement deviation that converts $\left\langle \widetilde{sl}\right\rangle$ to $\left\langle \widetilde{ml}\right\rangle$ is *not* self-enforcing since the common member country (i.e. country l) has an incentive to further deviate to become the hub country, taking the announcement of its complement as fixed. Third, country m always has no incentive to engage in any coalitional announcement deviations that replace $\left\langle \widetilde{sl}\right\rangle$ by $\langle sh\rangle$ or $\left\langle \widetilde{sl}\right\rangle$ by $\langle lh\rangle$. Third, note from the above discussion that the coalitional announcement deviation that replaces $\left\langle \widetilde{sl}\right\rangle$ by $\langle F\rangle$ is self-enforcing only when $\theta \leq \theta_l(F - \widetilde{sm})$. As a result, the announcement profile leading to $\left\langle \widetilde{sl}\right\rangle$ is a CPNE whenever $\theta \geq \theta_l(F - \widetilde{sm})$.

The following proposition summarizes our findings under the tariff discrimination scenario:

Proposition 4 *The equilibria of the tariff discrimination game of trade liberalization where FTA members have to practice free internal trade and the non-member do not have to abide by MFN and instead are free to tariff discriminate are as follows:*

(i) Free trade $\langle F\rangle$ is the equilibrium agreement when $\theta \leq \theta_l(F - \widetilde{sm})$.

(ii) Both asymmetric bilateral FTAs $\left\langle \widetilde{sl}\right\rangle$and $\langle \widetilde{sm}\rangle$ are stable when $\theta > \theta_l(F - \widetilde{sm})$

Insert Figure 2

Comparing the equilibrium outcome in the WTO-benchmark scenario (Proposition 2) and the equilibrium outcome in the tariff discrimination scenario (Proposition 3) yields the following result:

Proposition 5 *The comparison of the CPNE under the WTO-benchmark scenario and tariff discrimination scenario yields:*

(i) when $\theta \leq \theta_l(F - \widetilde{sm})$ holds, free trade is the equilibrium outcome under both tariff discrimination and the WTO-consistent benchmark;

(ii) when $\theta_l(F - \widetilde{sm}) < \theta \leq \theta_l(F - sm)$ holds, the WTO-consistent benchmark yields free trade whereas tariff discrimination yields bilateral FTAs $\widetilde{\langle sm \rangle}$ or $\widetilde{\langle sl \rangle}$ and

(iii) when $\theta > \theta_l(F - sm)$, free trade is out of reach under both scenarios but world welfare is lower under tariff discrimination.

Insert Figure 3

The above analysis provides strong support for the MFN principle. Not only does the MFN constraint make it easier to achieve global free trade, it also delivers a welfare-superior outcome when global free trade cannot be reached due to the high degree of asymmetry in the underlying economic environment. Yet, the above analysis cannot shed light on the effects of MFN when a bilateral agreement emerges because members deliberately exclude the third country since such an incentive on the part of members only arises when member countries can coordinate their external tariffs. Accordingly, in the second part of the paper, we discuss the case where the bilateral trade agreement is a customs union (CU) as opposed to an FTA.

5 Endogenous Customs Unions

Suppose the PTA under consideration is a CU under which member countries coordinate their external tariffs as opposed to an FTA. Similar to the FTA analysis above, we study CU formation in our WTO-consistent benchmark and then we investigate the tariff discrimination scenario.

First, consider the endogenous formation of CUs. As under the FTA game, at the first stage of the CU formation game each country announces the names of countries with whom it wants to form a CU. Country i's announcement is denoted by σ_i and its strategy set Ω_i consists of four possible announcements:

$$\Omega_i = \{\{\phi\}, \{j\}, \{k\}, \{F\}\} \tag{22}$$

where the announcement $\{\phi\}$ by country i is in favor of the status quo (or no trade liberalization); $\{j\}$ is in favor of a CU with only country j; $\{k\}$ is in favor of a CU with only country k; and $\{F\}$ is in favor of global free trade. As it is clear from the strategy set, CU formation is more rigid relative to FTA formation due to the common external tariff determination. Note that a hub and spoke type trading regime cannot arise under the CU game due to the fact that CU members coordinate their external tariffs. This announcement stage determines the global policy regime. Next, given the policy regime, countries impose their optimal external tariffs. Finally, given trade agreements and tariffs, international trade and consumption take place.

We obtain the following mapping between various announcements profiles: (i) no agreement $\langle\Phi\rangle$ prevails when no two announcements match or when everyone announces $\{\phi\}$; (ii) the CU $\langle ij\rangle$ is formed if countries i and j announce each other's name $\sigma_i = \{j\}$ and $\sigma_j = \{i\}$; (iii) free trade $\langle F\rangle$ obtains if $\sigma_i = \{F\}$ for all $i, j, k = s, m, l$.

We first state the following lemma that summarizes CU formation incentives under both the WTO-consistent benchmark and under tariff discrimination:

Lemma 3 *Regardless of whether the non-member country abides by MFN or not, the following results hold:*

(i) Each country prefers to form a bilateral CU with the larger importer relative to the smaller one.

(ii) The small importing country has no incentive to unilaterally or jointly deviate from any agreement.

(iii): A country is worse off as a non-member under a bilateral CU relative to no agreement, being a CU member and free trade

The above lemma informs us whether free trade obtains in equilibrium or not critically depends on the joint incentives of the medium and large importing countries to exclude the

small importing country. It turns out that this exclusion incentive arises only when the degree of endowment asymmetry is sufficiently large:

$$\begin{aligned} \Delta w_l(F - ml^u) &< 0 \text{ when } \theta > \theta_l(F - ml^u); \\ \Delta w_m(F - ml^u) &< 0 \text{ when } \theta > \theta_m(F - ml^u) \text{ and} \\ \theta_l(F - ml^u) &> \theta_m(F - ml^u) \end{aligned}$$

Here it is important to note that while the exclusion incentive go unexercised in the FTA formation game due to the relatively flexible nature of FTAs, such is not the case in the CU game. In the FTA game, if two countries (i and j) jointly exclude the third country from free trade by forming a bilateral FTA then each member has an incentive to sign an independent FTA with the excluded country thereby making itself a hub. The ability to act on this incentive acts as a deterrent for the other initially deviating country (say country j) and thus the initial joint deviation from free trade to a bilateral FTA does not occur. However, unlike the FTA game, no such deterrent exists under the CU game since a CU member cannot form an independent agreement with the excluded country without the consent of its CU partner. In other words, the joint announcement deviation leading to a deviation from free trade to $\langle ml^u \rangle$ is self-enforcing since neither country has an incentive to unilaterally deviate further. In fact, the joint deviation incentive of the large importing country is pivotal for the stability of free trade. As a result, the announcement profile Ω^F leading to free trade is a CPNE only when $\theta \leq \theta_l(F - ml^u)$.

Next, consider the announcement profiles leading to no agreement $\langle \Phi \rangle$. Starting from $\langle \Phi \rangle$, countries m and l have an incentive to jointly alter their announcements to form $\langle ml^u \rangle$. Since this deviation is self-enforcing, $\langle \Phi \rangle$ cannot arise in equilibrium. Similarly, based on parts (i) and (iii) of the Lemma 3, we directly argue that countries m and l have an incentive to jointly alter their announcement profiles so that $\langle sm^u \rangle$ and $\langle sl^u \rangle$ are replaced by $\langle ml^u \rangle$ and these deviations are self-enforcing since neither country m nor l has an incentive to unilaterally deviate further since doing so leads to $\langle \Phi \rangle$. As a result, the only remaining candidate for equilibrium is the announcement profile leading to $\langle ml^u \rangle$. Based on the above discussion, it is immediate that $\langle ml^u \rangle$ is a CPNE whenever $\theta \geq \theta_m(F - ml^u)$.

We summarize our main findings in the following proposition:

Proposition 6 *The equilibrium agreements in the CU game under the WTO-consistent scenario are as follows::*

(i) Free trade $\langle F \rangle$ is the unique CPNE when $\theta \leq \theta_m(F - ml^u)$.

(ii) Both free trade $\langle F \rangle$ and CU $\langle ml^u \rangle$ are stable when $\theta_m(F - ml^u) \leq \theta \leq \theta_l(F - ml^u)$.
(iii) Only CU $\langle ml^u \rangle$ is the unique CPNE if $\theta > \theta_l(F - ml^u)$.

Insert Figure 4

We are now ready to examine the implications of requiring MFN on the part of non-member country that has been excluded from the CU. As established earlier, countries will set a higher tariff on the country from whom it imports more from when free to tariff discriminate (Lemma 1). In the case of the stable CU $\langle ml^u \rangle$ in Proposition 3 above, this means that the excluded small country will set a higher tariff on the medium country relative to the large country. On one hand, faced with this tariff discrimination in the non-member's market, the medium country's incentive to exclude the small country weakens compared to the WTO benchmark scenario: $\theta_m(F - ml^u) < \theta_m(F - \widetilde{ml^u})$. On the other hand, since the large country faces a relative lower tariff under discrimination relative to MFN, this further strengthens its incentive to exclude the small country: $\theta_l(F - \widetilde{ml^u}) < \theta_l(F - ml^u)$. In fact, we show that these changes in the exclusion incentives of medium and large importing countries reverse the critical threshold rankings relative to the WTO consistent scenario: $\theta_l(F - \widetilde{ml^u}) < \theta_m(F - \widetilde{ml^u})$. Therefore, it is the joint deviation incentive of the medium importing country that determines the CPNE condition for free trade: the announcement profile Ω^F leading to free trade is a CPNE when $\theta \leq \theta_m(F - \widetilde{ml^u})$ holds. Similar to the WTO consistent scenario, we find that when $\theta \leq \theta_l(F - \widetilde{ml^u}) < \theta \leq \theta_m(F - \widetilde{ml^u})$ holds, both $\left\langle \widetilde{ml^u} \right\rangle$ and $\langle F \rangle$ are CPNE. Finally, when the degree of asymmetry is sufficiently large $\theta > \theta_m(F - \widetilde{ml^u})$ and global free trade fails to obtain due to exclusion incentives of the medium and large importing countries, $\left\langle \widetilde{ml^u} \right\rangle$ is the unique CPNE.

Combining the above results under tariff discrimination scenario with Proposition 6 has an interesting implication: when the threshold degrees of asymmetry are compared in the CU game under the WTO consistent and tariff discrimination scenarios, we obtain $\theta_l(F - ml^u) < \theta_m(F - \widetilde{ml^u})$. We can now state one of our main results:

Proposition 7 *For $\theta \leq \theta_l(F - ml^u)$ free trade is the equilibrium outcome for CUs under both tariff discrimination and the WTO-consistent scenarios. When $\theta_l(F - ml^u) < \theta \leq \theta_m(F - \widetilde{ml^u})$, tariff discrimination yields free trade whereas the WTO-consistent scenario yields $\left\langle \widetilde{ml^u} \right\rangle$. Finally, when $\theta > \theta_m(F - \widetilde{ml^u})$, free trade is out of reach under both scenarios and world welfare is higher under WTO consistent scenario in this tariff ridden world..*

Insert Figure 5

As noted before in the FTA game, by making tariff discrimination infeasible, the non-discrimination constraint of MFN reduces the benefit of being an outsider and thus lowers the unilateral incentive for opting out of trade liberalization. Therefore, in the FTA game, by increasing the likelihood of each country voluntarily choosing to enter into international trade agreements, the MFN principle can act as a catalyst for the cause of global free trade. However, when one country is *deliberately excluded* by the other two such as the case in the CU game, the MFN requirement for the non-member country can hinder the prospect of global free trade via strengthening the exclusion incentives. As a result, whether MFN complements Article XXIV in achieving global free trade depends on the nature of the trade agreements – i.e. whether it is an FTA or a CU. Finally, irrespective of the nature of the PTA in question, MFN adoption of the non-member country is world welfare improving when free trade fails to obtain either due to free riding incentive or exclusion incentive. Therefore, while whether existence of Article I with Article XXIV increases the likelihood of global free trade depends on the nature of PTAs, it is necessarily world-welfare improving in a tariff-ridden world.

6 Conclusion

In this paper, we have provided an analysis of the interaction between two core GATT rules governing trade liberalization at the WTO: Article 1 (MFN) that obligates members to adopt non-discriminatory trade policies toward one another and Article XXIV that lays down conditions that countries entering into PTAs are required to follow. To some degree, these two central GATT clauses clash with one another because although PTAs sanction discrimination against non-member countries, the MFN principle requires countries not participating in any PTAs themselves to refrain from discrimination amongst PTA participants. Motivated by this observation, we have developed a model that addresses two major questions. One, is there a case for allowing PTA non-members to tariff discriminate amongst PTA members? Two, does it matter whether a non-member contemplating such discrimination has *voluntarily* chosen to stay out of a PTA (knowing full-well that their non-participation will result in them facing discriminatory treatment) or finds itself *deliberately excluded* by PTA members? We show that holding the PTA non-member to MFN is desirable only when it chooses to voluntarily stay out of the trade agreement between the other two countries. Thus, there is no efficiency case for imposing MFN on countries

that find themselves excluded from a Customs Union that they wish to participate in but are barred from doing so by other member countries.

7 Appendix

In this Appendix, we detail all supporting calculations as well as proofs of lemmas and propositions.

7.1 Supporting calculations

First, we report the welfare levels as functions of an arbitrary tariff vectors. Using the welfare functions below and tariff levels reported in the text, we can calculate the formulae for optimum welfare levels under all possible regimes. Lemmas 1, 2, 3, and the various inequalities reported in the main text follow from a direct application of the relevant formulae.

7.1.1 Welfare levels

We report welfare levels for country i under a trade regime r as a function of an arbitrary tariff vector $\mathbf{t}(r)$ where $\mathbf{t}(r) = (t_{ij}(r), t_{ik}(r))$:

$$w_i(r) = \sum_z CS_i^z(r) + \sum_z PS_i^z(r) + TR_i(r)$$

where

$$\sum_z CS_i^z(r) = \frac{1}{2}\left[(\frac{e_j + e_k - t_{ij}(r) - t_{ik}(r)}{3})^2 + (\frac{e_i + e_k + 2t_{ji}(r) - t_{jk}(r)}{3})^2 + (\frac{e_i + e_j + 2t_{ki}(r) - t_{kj}(r)}{3})^2\right]$$

$$\sum_z PS_i^z(r) = \frac{e_i[6\alpha - 2e_i - e_j - e_k + t_{jk}(r) + t_{kj}(r) - 2t_{ji}(r) - 2t_{ki}(r)]}{3}$$

and

$$TR_i(r) = \frac{t_{ij}(r)[2e_j - e_k + t_{ik}(r) - 2t_{ij}(r)]}{3} + \frac{t_{ik}(r)[2e_k - e_j + t_{ij}(r) - 2t_{ik}(r)]}{3}.$$

7.2 Proofs of Lemmas and Propositions

Note that the proof of Lemma 1 is immediate from the optimal tariff discussion in the text.

Proof of Lemma 2

Using the above welfare formulae (as functions of an arbitrary tariff vectors) and plugging the above optimum tariffs into them, it is straightforward to show the following inequalities:

Part (i): $\Delta w_l(ml - sl) > 0$, $\Delta w_m(ml - sm) > 0$, and $\Delta w_s(sl - sm) > 0$ for all $1 \leq \theta \leq 5/4$.

Part (ii): $\Delta w_s(F - ml) > 0$ and $\Delta w_m(F - sl) > 0$ hold for all $1 \leq \theta \leq 5/4$ while $\Delta w_l(F - sm) < 0$ holds only when $\theta > \theta_l(F - sm) \cong 1.053$.

Part (iii): For the small country, $\Delta w_s(F - sm) > 0$ and $\Delta w_s(F - sl) > 0$ hold for all $1 \leq \theta \leq 5/4$. For the medium country, $\Delta w_m(F - ml) < 0$ only when $\theta > \theta_m(F - ml) \cong 1.1304$ while for the large country, $\Delta w_l(F - ml) < 0$ only when $\theta > \theta_l(F - ml) \cong 1.1023$.

Part (iv): $\Delta w_i(ih - \Phi) > 0$, $\Delta w_i(ih - F) > 0$ and $\Delta w_i(ih - ij) > 0$ for all $1 \leq \theta \leq 5/4$ and $i = s, m, l$.

Part (v): For the large country, $\Delta w_l(sh - sm) < 0$ and $\Delta w_l(mh - sm) < 0$ for all $1 \leq \theta \leq 5/4$. For the medium country, $\Delta w_m(lh - sl) < 0$ only when $\theta > \theta_m(lh - sl) \cong 1.204$ and $\Delta w_m(sh - sl) < 0$ only when $\theta > \theta_m(sh - sl) \cong 1.228$. For the small country, $\Delta w_s(lh - ml) > 0$ only when $\theta > \theta_s(lh - ml) \cong 1.0454$ and $\Delta w_s(mh - ml) > 0$ only when $\theta > \theta_s(lh - ml) \cong 1.0450$.

Proof of Proposition 1

Using the above welfare formulae (as functions of an arbitrary tariff vectors) and plugging the above optimum tariffs into them, it is straightforward to show the following:

- $\Delta w_s(sl - \Phi) > 0$, $\Delta w_m(sl - \Phi) > 0$, and $\Delta w_l(sl - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$.
- $\Delta w_s(sm - \Phi) > 0$, $\Delta w_m(sm - \Phi) > 0$, and $\Delta w_l(sm - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$.
- $\Delta w_l(ml - \Phi) > 0$ and $\Delta w_m(ml - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$ while $\Delta w_s(ml - \Phi) > 0$ holds only when $\theta < \theta_s(ml - \Phi) \cong 1.118$.

Proof of Proposition 2

Using the results from Lemma 2, the discussion in the main text and the following inequalities, it is straightforward to prove Proposition 2:

- $\Delta w_l(F - ml) < 0$ when $\theta > \theta_l(F - ml) \cong 1.102$;
- $\Delta w_m(F - ml) < 0$ when $\theta > \theta_m(F - ml) \cong 1.130$;
- $\Delta w_l(F - sh) < 0$ when $\theta > \theta_l(F - sh) \cong 1.213$;
- $\Delta w_l(F - mh) < 0$ when $\theta > \theta_l(F - mh) \cong 1.169$;
- $\Delta w_l(F - sm) < 0$ when $\theta > \theta_l(F - sm) \cong 1.0531$.

Proof of Proposition 3

Using the above welfare formulae (as functions of an arbitrary tariff vectors) and plugging the above optimum tariffs under discrimination into them, it is straightforward to show the following inequalities:

- $\Delta w_s(\widetilde{sl} - \Phi) > 0$; $\Delta w_l(\widetilde{sl} - \Phi) > 0$ and $\Delta w_m(\widetilde{sl} - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$.

- $\Delta w_s(\widetilde{sm} - \Phi) > 0$; $\Delta w_m(\widetilde{sm} - \Phi) > 0$ and $\Delta w_l(\widetilde{sm} - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$.

- $\Delta w_m(\widetilde{ml} - \Phi) > 0$, and $\Delta w_l(\widetilde{ml} - \Phi) > 0$ for all $1 \leq \theta \leq 5/4$ while $\Delta w_s(\widetilde{ml} - \Phi) > 0$ only when $\theta < \theta_s(\widetilde{ml} - \Phi) \cong 1.122$ where $\theta_s(ml - \Phi) \cong 1.118 < \theta_s(\widetilde{ml} - \Phi) \cong 1.122$.

Proof of Proposition 4

Using the results from Proposition 3, the discussion in the main text and the following inequalities, it is straightforward to prove Proposition 4:

- $\Delta w_l(F - \widetilde{ml}) < 0$ when $\theta > \theta_l(F - \widetilde{ml}) \cong 1.076$;

- $\Delta w_m(F - \widetilde{ml}) < 0$ when $\theta > \theta_m(F - \widetilde{ml}) \cong 1.219$;

- $\Delta w_l(F - sh) < 0$ when $\theta > \theta_l(F - sh) \cong 1.213$;

- $\Delta w_l(F - mh) < 0$ when $\theta > \theta_l(F - mh) \cong 1.169$;

- $\Delta w_l(F - \widetilde{sm}) < 0$ when $\theta > \theta_l(F - \widetilde{sm}) \cong 1.0526$ where $\theta_l(F - \widetilde{sm}) \cong 1.0526 < \theta_l(F - sm) \cong 1.0531$

Proof of Lemma 3

Using the above welfare formulae (as functions of an arbitrary tariff vectors) and plugging the above optimum tariffs under CUs into them, it is straightforward to show the following inequalities:

Part (i): Under the WTO-benchmark scenario: $\Delta w_l(ml^u - sl^u) > 0$, $\Delta w_m(ml^u - sm^u) > 0$, and $\Delta w_s(sl^u - sm^u) > 0$ for all $1 \leq \theta \leq 5/4$.

Under the tariff discrimination scenario: $\Delta w_l(\widetilde{ml^u} - \widetilde{sl^u}) > 0$, $\Delta w_m(\widetilde{ml^u} - \widetilde{sm^u}) > 0$, and $\Delta w_s(\widetilde{sl^u} - \widetilde{sm^u}) > 0$ for all $1 \leq \theta \leq 5/4$.

Part (ii): Under the WTO-benchmark scenario: For the small country, $\Delta w_s(sl^u - \Phi) > 0$, $\Delta w_s(sm^u - \Phi) > 0$, $\Delta w_s(F - sm^u) > 0$, $\Delta w_s(F - sl^u) > 0$, and $\Delta w_s(F - \Phi) > 0$ hold for all $1 \leq \theta \leq 5/4$.

Under the tariff discrimination scenario: For the small country, $\Delta w_s(\widetilde{sl^u} - \Phi) > 0$, $\Delta w_s(\widetilde{sm^u} - \Phi) > 0$, $\Delta w_s(F - \widetilde{sm^u}) > 0$, $\Delta w_s(F - \widetilde{sl^u}) > 0$, and $\Delta w_s(F - \Phi) > 0$ hold for all $1 \leq \theta \leq 5/4$.

Part (iii): Under the WTO-benchmark scenario: $\Delta w_s(ml^u - \Phi) < 0$, $\Delta w_m(sl^u - \Phi) < 0$, $\Delta w_l(sm^u - \Phi) < 0$, $\Delta w_s(ml^u - sm^u) < 0$, $\Delta w_s(ml^u - sl^u) < 0$, $\Delta w_m(sl^u - sm^u) < 0$, $\Delta w_m(sl^u - ml^u) < 0$, $\Delta w_l(sm^u - sl^u) < 0$, $\Delta w_l(sm^u - ml^u) < 0$, $\Delta w_s(F - ml^u) > 0$, $\Delta w_m(F - sl^u) > 0$, and $\Delta w_l(F - sm^u) > 0$ hold for all $1 \leq \theta \leq 5/4$.

Under the tariff discrimination scenario: $\Delta w_s(\widetilde{ml^u} - \Phi) < 0$, $\Delta w_m(\widetilde{sl^u} - \Phi) < 0$, $\Delta w_l(\widetilde{sm^u} - \Phi) < 0$, $\Delta w_s(\widetilde{ml^u} - \widetilde{sm^u}) < 0$, $\Delta w_s(\widetilde{ml^u} - \widetilde{sl^u}) < 0$, $\Delta w_m(\widetilde{sl^u} - \widetilde{sm^u}) < 0$, $\Delta w_m(\widetilde{sl^u} - \widetilde{ml^u}) < 0$, $\Delta w_l(\widetilde{sm^u} - \widetilde{sl^u}) < 0$, $\Delta w_l(\widetilde{sm^u} - \widetilde{ml^u}) < 0$, $\Delta w_s(F - \widetilde{ml^u}) > 0$, $\Delta w_m(F - \widetilde{sl^u}) > 0$, and $\Delta w_l(F - \widetilde{sm^u}) > 0$ hold for all $1 \leq \theta \leq 5/4$.

Proof of Proposition 6

Using Lemma 3 and the following inequalities, the proof is complete. We have $\Delta w_l(F - ml^u) < 0$ when $\theta > \theta_l(F - ml^u) \cong 1.0312$, $\Delta w_m(F - ml^u) < 0$ when $\theta > \theta_m(F - ml^u) \cong 1.0304$, and $\Delta w_m(F - sl^u) < 0$ hold for all $1 \leq \theta \leq 5/4$.

Proof of Proposition 7

Using Lemma 3 and the following inequalities, we can show the proof:

$\Delta w_l(F - \widetilde{ml^u}) < 0$ when $\theta > \theta_l(F - \widetilde{ml^u}) \cong 1.0213$, $\Delta w_m(F - \widetilde{ml^u}) < 0$ when $\theta > \theta_m(F - \widetilde{ml^u}) \cong 1.0541$, and $\Delta w_m(F - \widetilde{sl^u}) < 0$ hold for all $1 \leq \theta \leq 5/4$.

8 Figures

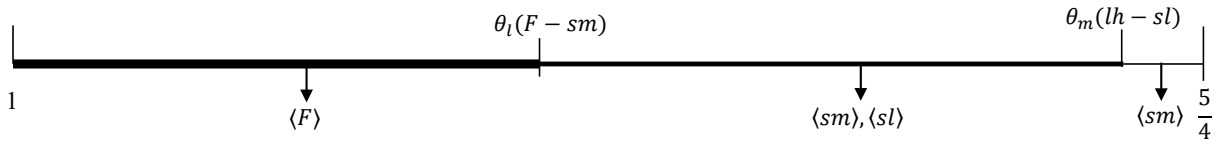

Figure 1: Equilibrium agreements under the benchmark WTO game with MFN (FTA)

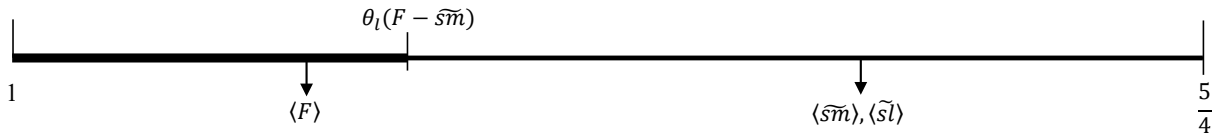

Figure 2: Equilibrium agreements with tariff discrimination (FTA)

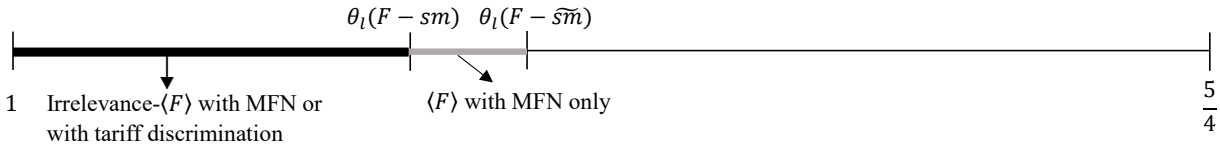

Figure 3: Free trade with MFN or with tariff discrimination (FTA)

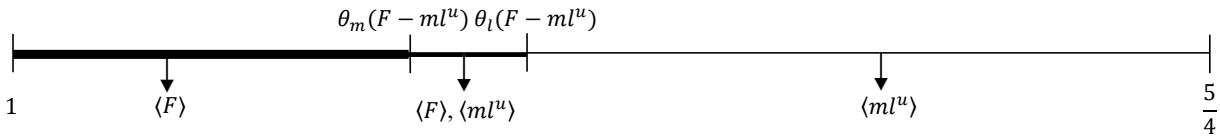

Figure 4: Equilibrium agreements under the benchmark WTO game (CU)

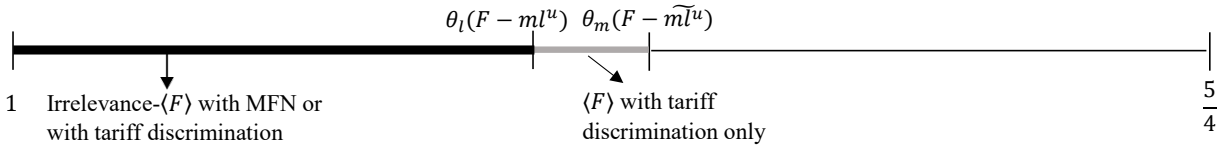

Figure 5: Free trade with MFN or with tariff discrimination (CU)

References

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

[2] Bagwell, Kyle and Robert. W. Staiger, 1999a. "Regionalism and Multilateral Tariff Cooperation." In John Piggott and Allan Woodland, eds, *International Trade Policy and the Pacific Rim*, London: MacMillan.

[3] Bagwell, Kyle and Robert W. Staiger, 1999b. "An Economic Theory of GATT." *American Economic Review* 89(1): 215-48.

[4] Bagwell, Kyle and Robert W. Staiger, 2010. "Backward Stealing and Forward Manipulation in the WTO." *Journal of International Economics* 82(1): 49-62.

[5] Bagwell, Kyle, Chad P. Bown, and Robert W. Staiger, 2016. "Is the WTO Passé?" *Journal of Economic Literature* 54(4): 1125–1231.

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

[7] Cadot, Olivier, Jaime De Melo, and Marcelo Olarreaga. 1999. "Regional integration and lobbying for tariffs against nonmembers." *International Economic Review* 40(3): 635-658.

[8] Caplin, Andrew, and Kala Krishna, 1998. "Tariffs and the Most Favored Nation Clause: A Game Theoretic Approach." *Seoul Journal of Economics* 1, 267-289.

[9] Choi, Jay P., 1995. "Optimal tariffs and the choice of technology: discriminatory tariffs vs. the most favored nation clause." *Journal of International Economics* 38: 143-160.

[10] Damuri, Yose Rizal, 2012. "How Preferential Are Preferential Trade Agreements? Analysis of Product Exclusions in PTAs." Graduate Institute Centre for Trade and Economic Integration Working Paper, 2012-3.

[11] Freund, Caroline 2000. "Multilateralism and the endogenous formation of preferential trade agreements." *Journal of International Economics* 52(2): 359-376.

[12] Freund, Caroline L., and Emanuel Ornelas, 2010. "Regional Trade Agreements." *Annual Review of Economics* 2: 139-166.

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

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

[15] Horn, Henrik and Petros Mavroidis, 2001. "Economic and legal aspects of the Most-Favored-Nation clause." *European Journal of Political Economy* 17(2) 233-279.

[16] Johnson, H.G., 1965. "An Economic Theory of Protectionism, Tariff Bargaining, and the Formation of Customs Unions." *The Journal of Political Economy* 72: 256-83.

[17] Kennan, John and Raymond Riezman, 1990. "Optimal Tariff Equilibria with Customs Unions." *Canadian Journal of Economics* 23(1): 70-83.

[18] Krishna, Pravin,1998. "Regionalism and Multilateralism: a Political Economy Approach." *Quarterly Journal of Economics* 113: 227–51.

[19] Liu, Xuepeng. 2010. "Testing Conflicting Political Economy Theories: Full-fledged versus Partial-scope Regional Trade Agreements." *Southern Economic Journal* 77(1): 78-103.

[20] Ludema, R.D. and Mayda, A.M., 2009. "Do countries Free Ride on MFN?" *Journal of international Economics* 77(2), pp.137-50.

[21] Ludema, R.D. and Mayda, A.M., 2013. "Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries." *The Quarterly Journal of Economics* 128(4): 1837-93.

[22] McCalman, Phil, 2002. "Multilateral trade negotiations and the most favored nation clause." *Journal of International Economics* 27: 151-76.

[23] Missios, Paul, Kamal Saggi and Halis M. Yildiz, 2016. "External Trade Diversion, Exclusion Incentives and the Nature of Preferential Trade Agreements." *Journal of International Economics* 99: 105-19.

[24] Olarreaga, M. and Soloaga, I., 1998. "Endogenous tariff formation: the case of Mercosur." *World Bank Economic Review* 12(2): 297-320.

[25] Ornelas, Emanuel, 2005a. "Trade Creating Free Trade Areas and the Undermining of Multilateralism." *European Economic Review* 49: 1717-35.

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

[27] Ornelas, Emanuel, 2007. "Exchanging market access at the outsiders' expense: the case of customs unions." *Canadian Journal of Economics* 40(1): 207-24.

[28] Ossa, Ralph, 2011. "A "new trade" theory of GATT/WTO negotiations. " *Journal of Political Economy*, 119(1): 122-152.

[29] Richardson, Martin, 1993. "Endogenous Protection and Trade Diversion." *Journal of International Economics* 34(3-4): 309-24.

[30] Richardson, Martin, 1995. "Tariff Revenue Competition in a Free Trade Area." *European Economic Review* 39(7): 1429-37.

[31] Saggi, Kamal, 2004. "Tariffs and the Most Favored Nation Clause." *Journal of International Economics* 63(2), 341-68.

[32] Saggi, Kamal and Halis M. Yildiz, 2009. "Optimal tariffs of preferential trade agreements and the tariff complementarity effect." *Indian Growth and Development Review* 4 (17).

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

[34] Saggi, Kamal, Woan Foong Wong and Halis M. Yildiz, 2019. "Should the WTO require free trade agreements to eliminate internal tariffs?" *Journal of International Economics*, 118: 316-330.

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

[36] Seidmann, Daniel, 2009. "Preferential Trading Arrangements as Strategic Positioning." *Journal of International Economics* 79: 143–59.

[37] Stoyanov, Andrey and Halis M. Yildiz, 2015. "Preferential versus Multilateral Trade Liberalization and the Role of Political Economy." *European Economic Review* 80: 140-164.

[38] Suwanprasert, Wisarut, 2018. "The Role of the Most Favored Nation Principle of the GATT/WTO in the New Trade Model " (Mimeo).

[39] Wong, Woan Foong, 2016. "The Free Rider Effect and Market Power in Trade Agreements." Working Paper.

[40] WTO. 2011. "World Trade Report 2011: The WTO and Preferential Trade Agreements: From Co-Existence to Coherence." Geneva: WTO.

[41] Yi, Sang-Seung, 1996. "Endogenous formation of customs unions under imperfect competition: open regionalism is good." *Journal of International Economics* 41(1): 153-77.