THE FARSIGHTED STABILITY OF GLOBAL TRADE POLICY ARRANGEMENTS

STEFAN BERENS, LASHA CHOCHUA†, AND GERALD WILLMANN‡

THE LATEST VERSION

Abstract. We study the stability of trade policy arrangements under different regulatory scenarios. Unlike previous papers, we consider an extensive set of trade policy arrangements and unlimited farsightedness of negotiating parties. In the standard three-country setup, global free trade (GFT) is uniquely stable under symmetric endowments. When two countries are small (large), the availability of Preferential Trade Agreements decreases (increases) the stability of GFT. With one type of PTA, strategic complementarity between Customs Unions (CUs) and Free Trade Agreements (FTAs) can ensure GFT stability. As asymmetry increases, GFT becomes unattainable, and the MFN regime is the only stable outcome without PTAs.

Keywords: Trade Policy Arrangements, Stability, Unlimited Farsightedness

JEL Classification: F13, F55

† IfW Kiel.

‡ Bielefeld University and IfW Kiel.

The following people supported and influenced this paper through various fruitful discussions: Herbert Dawid, James Lake, and Michael Chwe. Also, the authors are grateful for the generous financial support of the German Research Foundation (DFG) under the project WI 2038/6-1.

1. Introduction

Since the establishment of the General Agreement on Tariffs and Trade (GATT) in 1947, a growing number of countries have liberalized their trade policies via multilateral negotiations. To the present day, there have been eight rounds of multilateral trade negotiations, while the 9th round (started in Doha in 2001) has not been completed. At the same time, countries have also increasingly engaged in bilateral (or multi-party) talks to form preferential trade agreements, with 356 agreements[1] in force at the end of 2022.

The World Trade Organization (WTO, successor of the GATT since 1995) lays out the rules for international trade liberalization. Its Article I defines the Most-favored Nation (MFN) principle: any concession granted to one member shall be extended to all other members. In this paper, we will refer to trade policy arrangements that conform to the MFN principle as Multilateral Trade Agreements (MTAs). As an exception to the MFN principle, Article XXIV of the GATT/WTO allows countries to form regional or preferential trade agreements (PTAs). These take the form of Customs Unions (CUs) or Free Trade Agreements (FTAs) and allow countries to not extend the concessions granted within these argreements to other countries.[2]

Currently, 164 countries are members of the WTO,[3] and thus participate in the process of multilateral trade liberalization under the auspices of the WTO. Most of these countries are also engaged in preferential trade liberalization: Approximately twenty-five percent of them participate in more than ten preferential trade agreements (PTAs), while around forty percent are members of more than five PTAs.[4]

The objective of this study is to investigate the interplay of the simultaneous participation in the processes of multilateral and preferential trade liberalization. In this context, the effect of Article XXIV on the course of trade liberalization has long been a controversial topic, as evidenced by a raft of studies dealing with the issues involved. Building on this literature, we analyze the farsighted stability of a full set of trade policy arrangements under different scenarios: with PTA formation allowed as per current WTO rules, without either type of PTA (FTA or customs union), as well as with FTAs only, or only CUs. Our analysis sheds new light on the question of whether PTAs act as "building blocs" or "stumbling blocs" on the path toward global free trade (Bhagwati (1993)).

The existing literature on the stability of trade policy regimes usually considers only a limited set of trade agreements, or allows only limited farsightedness of the negotiating countries. While this allows for a cleaner description of the model and interpretation of its results, it raises the question of whether these restrictions may significantly influence the analysis, and to what extent such restrictive frameworks capture the essence of strategic interactions in reality. In our view, empirical observations favor a comprehensive choice set of trade policy

[1] Source: http://www.wto.org

[2] Article XXIV cautions, however, that these concessions shall not impinge on the trade relations between the member states and their other trading partners.

[3] These account for 96.4 percent of world trade, 96.7 percent of world GDP and 90.1 percent of the world population as of 2007 (Source: http://www.wto.org).

[4] Source: http://www.wto.org

regimes and full farsightedness. In fact, over past rounds of multilateral trade negotiations, many countries were simultaneously involved in both — multilateral as well as preferential — lines of trade liberalization.[5] Moreover, such trade negotiations are usually complex processes that significantly impact the countries' economies. Decision-makers therefore usually commission studies to quantify the effects and try to anticipate strategic actions by the trade partners involved.[6] Extending the analysis to a broad set of trade policy arrangements and taking into account the farsightedness of decision-makers is a crucial contribution of our work.

In this paper, we consider a comprehensive set of trade agreements comprising PTAs — both CUs, FTAs, and FTAHubs — as well as MTAs (multilateral agreements to lower tariffs under the MFN principle). When endogenizing the formation of trade agreements, we let each country rank these agreements based on a three-country two-good general equilibrium model of international trade.[7] We use the concept of 'largest consistent sets' (LCS) as stable outcomes — developed by Chwe (1994) — to capture the farsightedness of the negotiating parties.[8,9] Combining the welfare rankings of countries across regimes with the concept of 'largest consistent sets' (LCS) as stable outcomes allows us to determine the stability of the set of trade agreements over the entire parameter space. As a result, our work expands the set of trade agreements under consideration and, at the same time, extends the farsightedness of the negotiating parties in comparison to the existing literature. In fact, to the best of our knowledge, no other paper has considered a choice set as extensive as ours while at the same time including the concept of farsightedness.

When comparing the results from the scenario with a full range of trade agreements to an alternative setup without PTAs, our analysis shows that the effects of PTAs on trade liberalization depends on the size distribution of the countries in terms of their endowments.[10] As long as the countries are comparable in size, Global Free Trade (GFT) emerges as the unique stable outcome for both the existing and the alternative institutional arrangements (with vs. without PTAs). However, when two countries are considerably smaller than the third, a modified WTO arrangement preventing the formation of PTAs would facilitate the stability of GFT. This outcome arises from suppressing the small countries' exclusion incentive (when PTAs are available) vis-à-vis

[5] Maggi (2014) emphasizes the importance of an extensive set of trade policy constellations.

[6] Aumann and Myerson (1988) provide a (brief) description of the criticism against the use of limited farsightedness in general: 'When a player considers forming a link with another one, he does not simply ask himself whether he may expect to be better off with this link than without it, given the previously existing structure. Rather, he looks ahead and asks himself, "Suppose we form this new link, will other players be motivated to form further new links that were not worthwhile for them before? Where will it all lead? Is the end result good or bad for me?"'

[7] Our model is similar to the competing exporters model of Saggi and Yildiz (2010), which itself is a modification of the one in Bagwell and Staiger (1997). The modified version is also used in Saggi, Woodland, and Yildiz (2013). We use this model for our results to be comparable to the literature. Our framework can easily accommodate alternative models of trade.

[8] We thank Chew for sharing his algorithm which we have used as a blueprint to program the procedure that finds the LCS in our model.

[9] As a solution concept, LCS captures the notion of unlimited farsightedness in an introspective manner appropriate for static games. Coalitions make deviation decisions based on the comparison of the status quo with the final outcome of all possible deviations triggered by the initial deviation.

[10] Since endowments are exogenously given, they are used to measure the size of countries. The terms "small" and "large" refer to the characteristics of the supply side of the countries. As in Lake (2017), small countries can be viewed as the most "attractive" partners in the model, as they export relatively low quantities of goods and import relatively large quantities.

the large country. A modified WTO rule without PTAs, however, would actually reduce the extent to which GFT is stable if two countries are relatively larger than the third. A possible explanation for this result can be attributed to the free-riding incentive of the small country when PTAs are not in place. If the world is further away from symmetric endowments, full trade liberalization is not a stable outcome at all, and abolishing the option of PTAs can result in the most detrimental possible outcome from the perspective of overall world welfare, the non-cooperative MFN regime.

We also consider scenarios with only one type of PTA, either FTA or CU. In the scenario with only CUs (no FTAs allowed) the fundamental trade-off between the stability of GFT regimes with and without PTAs is still present as before. However, in a scenario with only FTAs (no CUs allowed) disallowing FTAs increases the area of GFT stability. Accordingly, if there are no free trade agreements, global free trade remains stable over a broader range of endowments. This suggests that free trade agreements in this scenario merely play the role of a "stumbling bloc". A further relevant comparison is between these two scenarios and the status quo, where both types of PTA are available. When switching off the possibility of forming CUs, even though there is a slight gain in the area where GFT is stable, we lose a bigger region where the GFT regime is no longer stable. By contrast, switching off FTAs does not result in any gain of stability for GFT, but rather results in a loss of stability in some areas of the parameter space. Consequently, the availability of FTAs and CUs in conjunction may be the most effective policy response to the free riding or exclusion incentives of small countries. To put it another way, under certain endowment constellations, strategic complementarity between CUs and FTAs ensures the stability of global free trade. In addition to its economic significance, this result also emphasizes the importance of taking into account the full range of possible trade policy arrangements when studying their stability. By focussing on FTAs or CUs only, such outcomes cannot be fully accounted for.

Our work is related to the literature on 'rules-to-make-rules' – as Maggi (2014) calls it – that tries to determine the role of institutions in the global trade liberalization process. In this field, there are some highly relevant studies. Critical among these contributions is the work of Kamal Saggi and Halis Yildiz, with several different co-authors. Saggi and Yildiz (2010) consider a three-country trade model where the degree and nature of trade liberalization – both bilateral as well as multilateral – are endogenously determined. Using Coalition-Proof Nash Equilibria (CPNE), the authors study the stability of FTAs and MTAs while varying the extent of asymmetry among countries with respect to their size. In a subsequent paper, Saggi, Woodland, and Yildiz (2013) study an analogous case, focusing on a combination of CUs and MTAs, leaving everything else fixed (in terms of their framework). By contrast, Missions, Saggi, and Yildiz (2016) analyze the effect of both forms of PTAs, i.e., CUs and FTAs, on attaining global free trade. However, the latter do not consider MTAs. In a sense, this research has hence been limited to studying '2 out of 3' types of trade agreements.

In Section 4 of the paper, we contrast our findings to those of Saggi, Woodland, and Yildiz (2013) wherever they are directly comparable. Our analysis reveals that certain predictions about stable outcomes based on the CPNE solution concept are not too persuasive. In contrast, the LCS solution concept provides a more consistent explanation of why different trade policy regimes should be considered as stable outcomes. For instance, to showcase the scope of differences in predictions, we refer to a scenario featuring two small and one larger country, where the endowment of the larger country varies. Saggi, Woodland, and Yildiz (2013) find that, for a significant portion of this interval, the only stable outcome is an MTA between the two small countries. Using the LCS concept, by contrast, there is again one stable outcome, but it is global free trade (GFT). We discuss the differences in results between the LCS and CPNE solution concepts, along with the advantages of our approach, in Section 4. Our findings suggest that the LCS solution concept is more robust and provides better predictions of the stability of different trade policy arrangements, especially in cases where the CPNE solution concept fails to provide convincing explanations.

Lake (2017), who employs a dynamic approach to understand whether FTAs facilitate or impede the path toward GFT, is another related paper (in terms of farsightedness). His approach involves using a three-country dynamic model in which a fixed protocol specifies the exact nature (and order) of negotiations for each period. The agreements that are formed are binding. In the end, using Markov Perfect Equilibrium in terms of pure strategies, Lake analyses the impact of country asymmetries on global trade liberalization. As discussed above, a similar question is addressed by Saggi and Yildiz (2010), who show that when certain conditions are met, bilateral free trade is necessary in order to achieve global free trade. Lake (2017) also demonstrates that FTAs are "strong building blocs" (global free trade is only possible in the presence of FTAs) if the discount factor is sufficiently small. But when the discount factor exceeds a certain threshold, they become "strong stumbling blocs". Contrary to these results, we find that FTAs never constitute 'building blocs' if only MTAs and FTAs are allowed. In our paper, FTAs can only be viewed as 'stumbling blocs' in a setup that resembles the one in Saggi and Yildiz (2010) and in Lake (2017).

In terms of the trade policy regimes considered, Lake (2019) is the most similar to our approach. Yet while he permits countries to choose between FTAs, CUs, and MTAs, he does not allow bilateral MFN liberalization, as we do in this paper. Additionally, the focus of his paper differs from ours. Lake studies the relative prevalence of FTAs versus CUs. On a related note, we must mention Seidmann (2009), whose research focuses on whether or not an FTA or CU is an equilibrium trade policy arrangement when transfers are available.

Having summarised the most related studies, we now consider the strand of work that deals primarily with analyzing the effect of FTAs. Goyal and Joshi (2006) consider several countries, each producing one good, studying different degrees of asymmetry across the countries. They employ the notion of *pairwise stability* proposed by Jackson and Wolinsky (1996) as a way of analyzing the stability of trade policy arrangements.

Furusawa and Konishi (2007) use similar methods but introduce a continuum of goods. Although they briefly discuss a setting with CUs in a separate section, the main thrust of their paper deals with FTAs. Another paper related to Goyal and Joshi (2006) is that of Zhang et al. (2013), who replace the concept of *pairwise stability* with *pairwise farsighted stability* that is due to Herings, Mauleon, and Vannetelbosch (2009). In this way, they compare myopia with farsightedness in an otherwise comparable framework. Connected to this is the paper by Zhang et al. (2014), which uses the work of Goyal and Joshi (2006) as a benchmark and analyzes the evolutionary effect of the number of countries in a dynamic framework featuring random perturbations. Whereas all these papers employ (different) network-theoretic concepts, Aghion et al. (2007), by contrast, features standard cooperative game theory. In their three-country model, a single country takes on the role of negotiation leader and decides to engage in sequential bilateral or single multilateral bargaining with the other countries.

Our study provides a comprehensive account of the effects of Article XXIV of the GATT/WTO on the liberalization process. We compare scenarios with and without preferential trade agreements (PTAs) to identify circumstances when PTAs act as "stumbling blocs" or "building blocs" on the path towards global free trade (GFT). In addition, we highlight the importance of strategic complementarity between Customs Unions (CUs) and Free Trade Agreements (FTAs) in ensuring the stability of global free trade in certain cases. We demonstrate that to counteract the free riding and exclusion incentives of small countries, FTAs and CUs should both be available simultaneously. As we move further away from symmetry, our analysis shows that Global Free Trade (GFT) is not attainable, and in the absence of PTAs, the non-cooperative most-favored nation (MFN) regime will be the only stable outcome. Our findings underscore the importance of PTAs in facilitating trade and promoting global welfare. Overall, we highlight the importance of considering strategic behavior and complementarity between different trade policy arrangements, such as CUs and FTAs, to promote global free trade and welfare.

Finally, our paper incorporates the stability concept introduced by Chwe (1994). This concept serves as a response to address certain criticisms of the von Neumann-Morgenstern stable set solution.[11] The application of Chwe's approach accomplishes two key objectives in our study: first, to enable participants to consider a wide range of future possible deviations; and, second, to avoid the emptiness of the stable set that plagues other (more) restrictive solution concepts.[12] This stability concept bears close resemblance to the concept proposed by Herings, Mauleon, and Vannetelbosch (HMV) (2009) and its subsequent extension in HMV (2014).

The structure of the paper is as follows. In Section 2, we introduce the model and present the solution concept. Section 3 provides the main analysis and describes the findings of our study. In Section 4, we compare our results to those of the most relevant papers in the literature. Finally, Section 5 offers concluding remarks.

[11] Consult von Neumann and Morgenstern (1944) for a description of this (solution) concept and Harsanyi (1974) for its criticism.

[12] It is also resistant to the criticism of Ray and Vohra (2015) about the sovereignty of coalitions as their main issues concerned with feasibility and distribution do not apply to our framework. Furthermore, their specific criticism about the explanatory power of Chwe's approach only applies to transferable utility games.

2. Model

This section presents the trade policy arrangements we consider and introduces the transition graphs that show how countries can move from one policy regime to another, individually or in coalitions. Subsequently, we provide the necessary theoretical and intuitive details of the stability notion we use, including the solution concept. Finally, we present the underlying trade model that determines countries' preferences over trade policy regimes and parametrize the model.

2.1. Trade Policy Arrangements.

As is standard in the literature, we consider a three-country world. Let $N = \{a, b, c\}$ denote the set of three countries. These countries trade with one another, and each country can (possibly) impose tariffs on imported goods following the rules of the GATT. Let t_{ij} denote the tariff (vector) imposed by country i on imports from country j, where $i, j \in N$ are distinct. The decision-makers in each country maximize national welfare. We denote by W_i the welfare of country $i \in N$. This basic setup allows us to study all possible trade policy constellations under articles I and XXIV of the GATT. In subsection 2.4, we provide further details on the underlying trade model, which then allows us to determine the tariffs chosen and the resulting welfare levels, conditional on trade policy regime, which in turn determine countries' preferences over these regimes.

All trade policy arrangements we consider are of the following four types: most favored nation (MFN), customs union (CU), free trade area (FTA), and multilateral trade agreement (MTA). Each type (except for MFN) induces combinations of insiders and outsiders. For a CU, (single) FTA, and MTA, we have three combinations each of two members, plus one comprising the full set of countries. Note that we take Global Free Trade to come in three different forms: a CU, FTA, or MTA of three members.[13] In addition, there is the case of two over-lapping FTAs in a hub-and-spoke structure — adding another three regimes. In total, we thus have 16 different global trade policy regimes and define the set of global trade policy constellations as $X \equiv \{MFN, CU(i,j), CU(i,k), CU(j,k), CUGFT, FTA(i,j), FTA(i,k), FTA(j,k), FTAHub(i), FTAHub(j), FTAHub(k), FTAGFT, MTA(i,j), MTA(i,k), MTA(j,k), MTAGFT\}$, where $i, j, k \in N$.

Under each of these trade agreements, tariffs are bounded by zero from below, and from above by the MFN-tariff, which we discuss in more detail in Appendix B.2. The national policymaker in country i thus chooses $t_{ij} \in T_i \equiv [0, t_i^{MFN}]$ for $\forall j$.[14] Conditional on trade agreement, the choice of tariffs is restricted further, as we discuss now.

In the baseline case, which we denote by MFN, countries do not liberalize their trade relations but satisfy the non-discrimination principle. Each country i unilaterally chooses its (optimal) tariffs, solving $\max_{(t_{ij}, t_{ik}) \in T_i^{MFN}} W_i$

[13]While their welfare level is equal, their position in the network (cf. Section 2.2) need not be, and for our concept of stability, it is essential which group of countries can create or destroy specific trade agreements. Where applicable, we group all three and refer to them as 'GFT'.

[14]Note that the upper bound is uniform across trading partners by the very nature of MFN.

with $T_i^{MFN} = \{(t_{ij}, t_{ik}) \in \mathbb{R}^2_{\geq 0} \,|\, t_{ij} = t_{ik}\}$. Note, that in this reference scenario each tariff is chosen from $\mathbb{R}_{\geq 0}$ instead of T_i.

When countries i and j form a customs union $CU(i,j)$ together, each of them removes any trade restriction on the partner country, and they jointly impose an optimal tariff on country k. Their optimization problem amounts to $\max_{(t_{ij},t_{ik})\in T_i^{CU},(t_{ji},t_{jk})\in T_j^{CU}} W_i + W_j$ with $T_i^{CU} = \{(t_{ij}, t_{ik}) \in T_i^2 \,|\, t_{ij} = 0 \text{ and } t_{ik} = t_{jk}\}$ and analogously for T_j^{CU}. Country k, finally, simply follows and applies the MFN principle (as before). Once all three countries enter a single CU together, the (common) optimisation problem is trivial because internal tariffs are restricted to zero. We denote this scenario by CUGFT.

In case countries i and j form an $FTA(i,j)$, each of them again removes any trade restriction on the other country, and unilaterally imposes an optimal tariff on the outsider country k. The (representative) optimisation problem of country i (and j) is $\max_{(t_{ij},t_{ik})\in T_i^{FTA}} W_i$ with $T_i^{FTA} = \{(t_{ij}, t_{ik}) \in T_i^2 \,|\, t_{ij} = 0\}$. The optimization problem of country k is identical to that of the third country in the case of a CU. Furthermore, if country i forms an FTA each with countries j and k, we call this $FTAHub(i)$, then both tariffs of country i are set to zero by nature of its trade relationship with both partner countries. Each of the other two countries operates as before: Country j (k analogous) solves $\max_{(t_{ji},t_{jk})\in T_j^{FTA}} W_j$ where $T_j^{FTA} = \{(t_{ji}, t_{jk}) \in T_i^2 \,|\, t_{ji} = 0\}$. Note that in terms of the decision problem, it does not matter for a country whether its partner also forms a trade agreement with the other country. Finally, if all three countries form an FTA, denoted by FTAGFT, the optimization problem is identical to the case of CUGFT but differs in terms of structure and network position (cf. Section 2.2).

We turn finally to the case where countries i and j form a multilateral trade agreement $MTA(i,j)$. In this case, they jointly change their tariffs vis-à-vis each other, and also for the third country. The corresponding optimization problem is $\max_{(t_{ij},t_{ik})\in T_i^{MTA},(t_{ji},t_{jk})\in T_j^{MTA}} W_i + W_j$ with $T_i^{MTA} = \{(t_{ij}, t_{ik}) \in T_i^2 \,|\, t_{ij} = t_{ik}\}$ and analogously for T_j. As before, the optimization problem of country k is identical to that of the outsider country in the case of a CU. When all three countries enter a single MTA together, which we denote by MTAGFT, the optimization problem is identical to the case of CUGFT and FTAGFT, but the regime differs in terms of network position (cf. Section 2.2).

2.2. **Transition Graphs.** We now consider transition possibilities from one global trade policy constellation to another. Countries can implement trade policy arrangements, or leave an existing arrangement, depending on whether they act individually or in a coalition with other countries. We represent these possibilities using directed graphs, with trade agreements as vertices and the transition between them as directed edges.

To start with, consider the case of a single country $i \in N$ (this being a coaltion of one), with $j, k \in N \setminus \{i\}$, $j \neq k$ denoting the other two countries. The digraph for this case is provided in Figure 1, leaving aside loops at every vertex for illustrative clarity. As we can see from the diagram, the baseline MFN regime is reachable from a number of different global trade policy regimes, which country i can decide to quit, but not from the trinity of

GFT, i.e., CUGFT, FTAGFT, and MTAGFT, because quitting those would leave a two-country arrangement between j and k in place (or possibly an FTA hub-structure, if i were to quit only one FTA). The three variants of GFT thus form separate groups of connected trade agreements, and hence the overall transition graph (in Figure 1, abstracting from loops), consists of four sub-graphs.

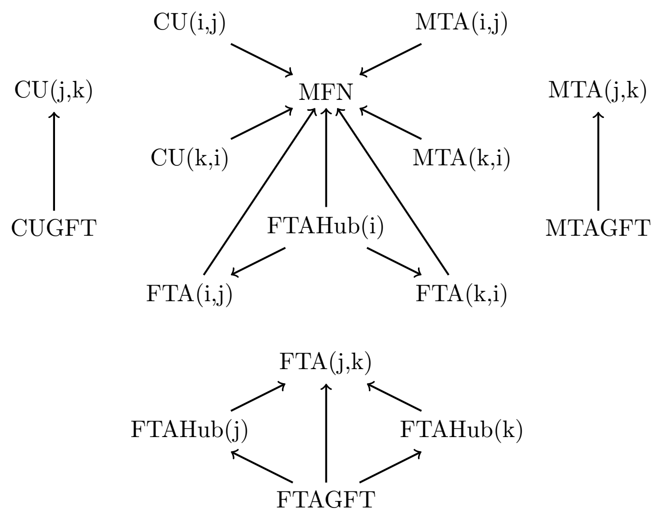

FIGURE 1. The transition graph for coalition $\{i\}$, $i \in N$.

Next, consider the transition graph for a coalition of two countries $i, j \in N$, where $i \neq j$ and $k \in N \setminus \{i, j\}$ denotes the other country. The transition graph for this case is provided in Figure 2, again leaving aside loops. As we see here, the four regimes MFN, CU(i,j), FTA(i,j), and MTA(i,j) are all interconnected. This is because the coalition of i and j can move from and to any two-country agreement between themselves and MFN at will. In addition, any regime connected to one of the four is automatically connected to all four of them. This group of four represents a complete directed sub-graph, which we illustrate as a dotted (super) node in the figure. In contrast to the previous case (compare Figure 1), the coalition of i and j can move from GFT to MFN or any the two-country arrangement between themselves (the super node). They can also transition from any form of GFT to a corresponding two-country arrangement between one of them and the outside country, if only one of them decides to quit, plus from FTAGFT to any FTA hub-and-spoke structure.

Finally, consider the coalition of (all) three countries. Together, these three can implement any trade policy constellation. Therefore, the corresponding directed graph for the case of the grand coalition is a complete directed graph with loops (for short, a complete loop-digraph).

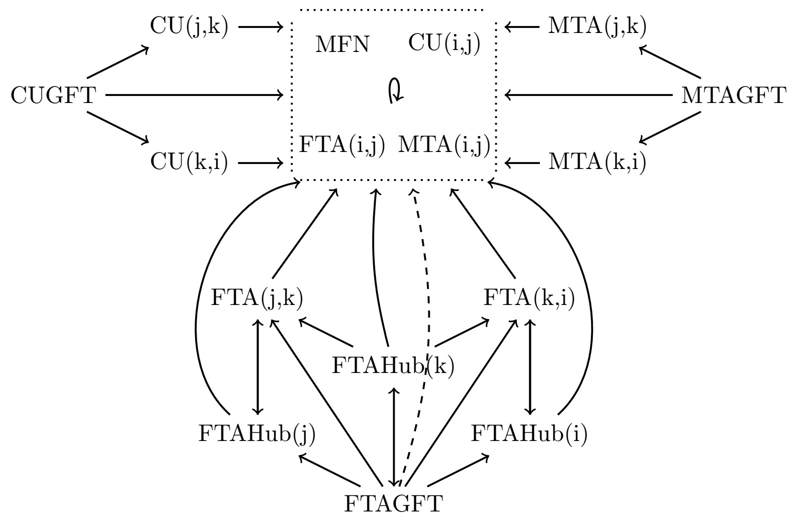

FIGURE 2. The transition graph for coalition $\{i,j\}$, $i,j \in N$, $i \neq j$.

2.3. **Stability Concept.** In terms of the stability concept, we use the approach of Chwe (1994), as he combines farsightedness with deviations by a broader, more realistic set of coalitions.[15]

Consider the tuple $\Gamma = \big(N, X, \{\prec_i\}_{i\in N}, \{\rightarrow_S\}_{S\subseteq N, S\neq\emptyset}\big)$ that lists the set of countries or players, the set of global trade policy constellations, the preferences of players over these regimes, and the effectiveness relation of possible transition paths. To understand the meaning of the latter, let $x \in X$ be the status quo trade agreement at the start. Next, each coalition $S \subseteq N$, $S \neq \emptyset$ (including one-country coalitions) is able to make $y \in X$ the new status quo as long as $x \rightarrow_S y$. Continue with y as the new status quo. Suppose a status quo $z \in X$ is reached without any coalition moving away. In that case, the state is realized, and each country receives its corresponding welfare.[16] As a consequence, any coalition only favors following through on their ability to move from $x \rightarrow_S y$, if it prefers the final welfare over the current one, $x \prec_S z$. Formally, this comparison of states by (chains of) coalitions is captured in the definition of direct and indirect dominance:

Definition 1 (Dominance)**.** *Let $x_1, x_2 \in X$. Then,*

i) *x_1 is directly dominated by x_2, write $x_1 < x_2$, if there exists $S \subseteq N$, $S \neq \emptyset$, such that $x_1 \rightarrow_S x_2$ and $x_1 \prec_S x_2$.*

ii) *x_1 is indirectly dominated by x_2, write $x_1 \ll x_2$, if there exist sequences $y_0, y_1, \dots, y_m \in X$ (with $y_0 = x_1$ and $y_m = x_2$) and $S_0, S_1, \dots, S_{m-1} \subseteq N$, such that $S_i \neq \emptyset$, $y_i \rightarrow_{S_i} y_{i+1}$, and $y_i \prec_{S_i} y_m$ for $i = 0, 1, \dots, m-1$.*

Note, that direct dominance implies indirect dominance, i.e. if $x_1 < x_2$ for some $x_1, x_2 \in X$, then automatically $x_1 \ll x_2$.

[15] Consult the paper of Chwe (1994) for proofs of the propositions that are presented here.

[16] Technically, the model is without any true sense of time. Any start (or end) and any sequence of actions should be interpreted as a thought experiment. Furthermore, a path created in this fashion is generally not unique.

Using this definition, we introduce the concept of the 'consistent set', a (sub-)set that exhibits internal stability in the sense of a lack of incentives to deviate:

Definition 2 (Consistent Set)**.** *A set $Y \subseteq X$ is consistent if $y \in Y$ if and only if for all $x \in X$ and all $S \subseteq N$, $S \neq \emptyset$, with $y \rightarrow_S x$ there exists $z \in Y$ where $x = z$ or $x \ll z$ such that $y \not\prec_S z$.*

In general, a consistent set need not be unique. Still, the following proposition allows us to focus on the unique 'largest consistent set', i.e., the (consistent) set that contains all consistent sets:

Proposition 1. *There exists a unique $Y \subseteq X$ such that Y is consistent and $Y' \subseteq X$ consistent implies $Y' \subseteq Y$. The set Y is called the largest consistent set or simply LCS.*

Similar to the internal stability captured by the definition of consistent sets, external stability is captured by an incentive to gravitate toward the consistent set:

Definition 3 (External Stability)**.** *Let $Y \subseteq$ be the largest consistent set. Then, it satisfies the external stability condition if for all $x \in X \setminus Y$ there exists $y \in Y$ such that $x \ll y$.*

The following result characterises one setting of relevance in which this condition is satisfied:

Proposition 2. *Let X be finite and the underlying preferences irreflexive. Then, the LCS is non-empty and satisfies the external stability property.*

Importantly, this result applies to our setting, so the (unique) LCS is non-empty and satisfies external (as well as internal) stability.[17]

The largest consistent set is going to be the focus of our analysis. Any trade agreement is considered '(potentially) stable' if it is in the LCS and 'unstable' otherwise. Note that the nomenclature is a tribute to the fact that the LCS as a stability concept is 'weak: not so good at picking out, but ruling out with confidence', because ultimately, it 'does not try to say what will happen but what can possibly happen' (Chwe (1994)).

2.4. **Underlying Trade Model.** To study the stability of different global trade policy arrangements, we use a three-country, general equilibrium trade model. We follow the literature by using the model employed by Saggi and Yildiz (2010), which features competition in exports. We want to emphasize that our approach allows for the use of any type of underlying trade model. Our choice is guided solely by the widespread use of the competing exporters model in the literature, so that the use of this model facilitates the comparison of results.

[17]First of all, applying Proposition 1 to our model is trivial because it is stated without any (additional) requirements on the involved objects. Furthermore, the application of Proposition 2 is straightforward as well: First, the set of outcomes X is clearly finite in our setting as we are only considering a finite number of different trade agreements. Second, any strict preference is automatically irreflexive, and our preferences are induced by strict welfare comparisons. Thus, while the definition of the (largest) consistent set in general only guarantees internal stability, our setting actually implies external stability as well.

Recall that $N = \{a, b, c\}$ denotes the set of countries. We denote by $G = \{A, B, C\}$ three non-numéraire goods, so that each country i is endowed with zero units of good I (the corresponding capital letter), and e_i units of the other two goods. Country i will end up importing good I and exporting goods J and K, with $J, K \neq I$. Each good will thus be exported by two countries, e.g., good I will be exported by countries j and k, which has given the model its name, 'competing exporters model'. In order to guarantee the 'competing exporters'-structure, we need to impose a condition on the degree of asymmetry in terms of the respective endowments:

$$\frac{3}{5}\max\{e_j, e_k\} \leq e_i \leq \frac{5}{3}\min\{e_j, e_k\} \qquad \forall\, i, j, k \in N$$

This condition is sufficient to ensure that exports remain non-negative.

On the demand side, let preferences of individuals in each country be identical. The demand for any non-numéraire good $L \in G$ in country $i \in N$ is given by $d(p_i^L) = \alpha - p_i^L$ with p_i^L the price of good L in country i and α the (universal) reservation price.[18] As pointed out before, each country can impose tariffs on imports. Let t_{ij} denote the tariff imposed by country i on the import of good I from country j. The prices and tariffs of good $I \in G$ across countries are linked via the following no-arbitrage condition:

$$p_i^I = p_j^I + t_{ij} = p_k^I + t_{ik}\ , \tag{1}$$

where $i, j, k \in N$ are pairwise distinct.

In the model at hand, prices together with the corresponding endowments are the only determinants of imports and exports. In particular, the level of imports m_i^I of good I into country i is completely determined by the above demand function (depending on the price): $m_i^I = d(p_i^I) = \alpha - p_i^I$. In turn, the exports x_j^I of good I from country j are the combination of the demand function (or prices) and the corresponding endowment, $x_j^I = e_j - d(p_j^I) = e_j + p_j^I - \alpha$. Market-clearing for any good I requires that country i's imports equal the exports of countries j and k (again $i, j, k \in N$ pairwise distinct) combined:

$$m_i^I = x_j^I + x_k^I \tag{2}$$

Ultimately, the objective function of the benevolent policymaker in country i is its national welfare[19], denoted W_i, which includes consumer surplus (CS), producer surplus (PS), and tariff revenue (TR):

$$W_i = \sum_{L \in G} CS_i^L + \sum_{L \in G \setminus \{I\}} PS_i^L + TR_i$$

[18] The demand function can be derived from a utility function that is additively separable and quadratic in each non-numéraire good.

[19] Depending on the type of trade agreement entered, the (joint) objective function may include the welfare of other countries as well. See Section 2.1 for the details.

Using the no-arbitrage (1) and market-clearing (2) conditions, we can compute the equilibrium prices:

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

Given equilibrium prices, we can deduce imports, exports, and the welfare of each country up to the value of the tariffs (Appendix B.1). The maximization of welfare with respect to tariffs will then be constrained by the trade agreement under consideration (see Section 2.1). The full equilibrium of our model is then computed as follows: Fix a trade agreement and thereby the restrictions on tariffs. Compute the best-response functions for each country (with respect to the tariffs) and determine the optimal choices. While Section 2.1 contains all information on the trade agreements that is necessary to compute the equilibria, the actual results are presented in Appendix B.2. An overview of the (resulting) overall welfare is provided in Appendix B.3.

2.5. **Algorithm and Parameters.** The contribution in this paper of considering an extensive set of trade agreements and unlimited farsightedness comes at the cost of a relatively complex computational problem. We solve this problem numerically by means of an algorithm, the pseudocode of which we provide in Appendix A.[20] To solve the model numerically using the algorithm, we need to specify and discretize the parameter space.

Recall that the endowments have to satisfy $\frac{3}{5}\max\{e_j, e_k\} \leq e_i \leq \frac{5}{3}\min\{e_j, e_k\}$ for all $i, j, k \in N$ in order to guarantee the 'competing exporters'-structure (cf. Section 2.4). Without loss of generality, let us normalize one country's endowment to one, namely $e_b = 1$. Consequently, for $i, j \in N \setminus \{b\}$: $e_{\min} \equiv \frac{3}{5} \leq \frac{3}{5}\max\{1, e_j\} \leq e_i$ and $e_i \leq \frac{5}{3}\min\{1, e_j\} \leq \frac{5}{3} \equiv e_{\max}$. The resulting parameter space is depicted by the hexagon in Figure 3. Given that the three countries are interchangeable, we can split the hexagon into six right-angled triangles, which are mirror images of one another (in terms of relative endowments). Without loss of generality, we focus on one of them, namely the triangle depicted as shaded in the figure. We cover this triangle of the parameter space with a fine grid for the actual computation.[21]

In addition to endowments, we also need to specify the choke price *alpha* in the demand function. To obtain plausible results — that is, positive prices — the parameter α needs to be chosen above a minimal value for each tuple of endowments, $\alpha_{\min}(e_a, e_b, e_c)$. Above this minimal value, results remain unchanged.[22] Taking the maximum over all these minimal values, $\alpha_{\max\min} = \max_{e_a,e_b,e_c}\{\alpha_{\min}(e_a, e_b, e_c)\}$, adding an epsilon, $\alpha = \alpha_{\max\min} + \epsilon$, and using it for all endowments combinations ensures that the results are plausible and comparable.[23]

[20] The authors are grateful to Michael Chwe for the provision of an exemplary algorithm.

[21] The distance is set to 0.0013360053440215, resulting in 500 points per dimension.

[22] The parameter α enters the welfare of country i as $2\alpha e_i$ (see Appendix B.1). Any changes above the minimal value leave the relative welfare levels and therefore the rankings unaffected.

[23] In our computation ϵ is fixed at 0.01, which yields $\alpha = 1.399$.

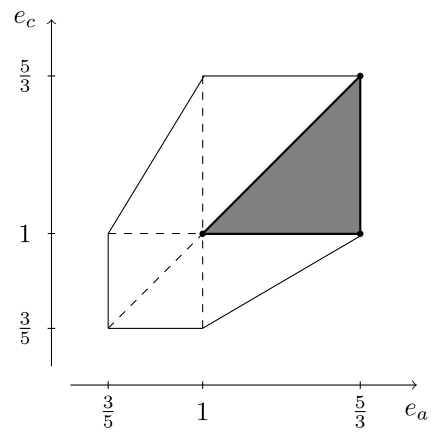

FIGURE 3. The parameter space of the endowments with $e_b = 1$

3. Analysis

We now turn to the analysis of stable outcomes among the global trade policy regimes that we consider. That is, for every endowment vector, we determine which regimes make up the largest consistent set. Figure 4 depicts the parameter space of endowments under consideration — it is the (marked) triangle from before. Note that different endowment vectors represent varying size combinations of countries.

For expositional purposes, we start by studying the three corners of the triangle, then turn to the connecting intervals or edges, and finally, consider the entire triangle by examining the whole interior. Clearly, the corners are the extrema of the edges, and the edges, in turn, the bounds of the interior. So the second and the third step contain the preceding results as special cases, and we can think of each subsequent step as a linear combination of the endowment combinations considered previously.

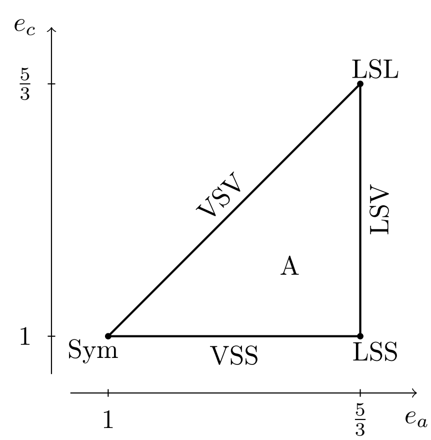

FIGURE 4. Overview of the different points, intervals and areas of interest depending on the (partially normalised) endowment tuple

For all allowable endowment combinations, we will analyze two scenarios: The first scenario corresponds to the current institutional setup of the WTO. The constitutional status-quo includes PTA regimes, as article

XXIV provides for. The second scenario that we study is a hypothetical alternative constitutional arrangement without article XXIV, i.e., the choice set of global trade policy regimes does not include any form of PTA (specifically, neither CUs nor FTAs).

3.1. **Corners of the triangle.** We start our analysis by considering the three corners of the triangle. First, thorough attention will be given to the symmetric case to demonstrate the working of each component of the model. Subsequently, we focus on results for the other two corners and provide a general overview.[24]

In the symmetric corner, denoted by Sym in Figure 4, all three countries have identical endowments: $e_a = e_b = e_c = 1 = e_{\min}$. As the countries do not differ from one another in this case, the only thing that matters for welfare is whether a country is an insider or an outsider in a specific trade agreement. In the following, we present the ranking of preferences from the perspective of country a (which represents the preferences of the other countries as well). For fixed $i, j \in N \setminus \{a\}$ with $i \neq j$, we obtain the following preference ranking:

$$\begin{aligned} CU(i,j) \prec_a MFN \prec_a MTA(a,i) \prec_a FTAHub(i) \prec_a FTA(i,j) \prec_a FTA(a,i) \\ \prec_a CU(a,i) \prec_a MTA(i,j) \prec_a GFT \prec_a FTAHub(a) \end{aligned}$$

The case where the two other countries i and j form a customs union is the least favorable trade constellation for country a. As an outsider to the $CU(i,j)$, country a faces the second-highest tariffs (with MFN-tariffs the highest), while the CU members i and j abolish the tariffs between themselves. The exports of country a to the other countries, i and j, are the lowest under CU(i,j) compared to all alternative trade agreements. The same applies to total imports. In other words, the 'trade diversion' effect is strongest for country a when we have CU(i,j).

In general, the MFN regime favors country a as compared to CU(i,j). The tariff revenue remains the same, while the consumer surplus is lower and the producer surplus higher — the increase in the latter offsetting the decrease of the former. The MFN regime mitigates the 'trade diversion' effect present in the case of CU(i,j) due to increased export values of country a.

Within the set of bilateral trade agreements where country a is an insider, the MTAs result in the lowest welfare (for country a). MTA(a,i) generates higher welfare for country a in comparison to the MFN regime due to increased consumer and producer surplus. The FTAHub(i) constellations result in further gains in welfare for country a through higher export values and producer surplus, even though tariff revenue and consumer surplus are lower under FTAHub(i) compared to MTA(a,i).

However, country a does not have an incentive to remain in this constellation. The unilateral deviation from FTAHub(i) to FTA(i,j)— that is, if a quits its FTA with the hub country — comes with a decrease in consumer and producer surplus but enough of an increase in tariff revenue to ultimately ensure higher welfare

[24]We provide additional proofs for two, non-symmetric corners in Appendix C.1.

under FTA(i,j). Yet, among (single) FTAs, being an outsider is less desirable than being an insider for any of the three countries, as the drop in tariff revenue is offset by an expansion of consumer and producer surplus, resulting in higher welfare for the country a in the case of FTA(a,i) compared to FTA(i,j). As an insider, country a prefers CU(a,i) over FTA(a,i), though. Despite the decline in the consumer surplus, welfare goes up due to an expansion of tariff revenue and producer surplus.

The formation of MTA(i,j) guarantees the highest welfare for the country a compared to any other bilateral trade agreement. The driving factor is the MFN principle, which implies that in case of MTA(i,j) the insiders need to apply the same tariff to their fellow member and to the outsider. Country a free-rides on this characteristic of the MTA regime, and in addition, attains the highest possible tariff revenue in this case.

Each country obtains the second-highest welfare level when the world reaches global free trade (that is, the trinity of MTA-, FTA-, and CU-GFT). Under full trade liberalization, the producer surplus is at its second-highest among all trade agreements (effectively driving the ranking). It is only surpassed by that of FTAHub(a). This constellation brings about the highest possible welfare for the country a. Notice, however, that this hub-n-spoke trade agreement disproportionally favors the hub country over the other two.

The countries' preference rankings, which we have just derived for one particular endowment combination, provide the basis for computing the largest consistent set (LCS). In the case at hand, the trinity of global free trade regimes are ranked second-best by each country, while the respective first-best options, hub-n-spoke FTA structures, are ranked considerably lower by spoke countries. The trinity of global free trade thus seems the likely equilibrium outcome. The following proposition and its proof confirm this:

Proposition 3. *Under symmetry and with the current institutional arrangement of the WTO, the LCS contains three elements: CUGFT, FTAGFT, and MTAGFT. In other words, (the trinity of) global free trade is the unique stable outcome.*

Proof. Based on the definition of indirect dominance and the transition graphs (cf. Sections 2.3 and 2.2), the preference rankings from above allow us to derive the indirect dominance matrix. If the entry in the matrix is equal to one (resp. zero), then the trade arrangement corresponding to the row of the entry is (respectively is not) indirectly dominated by the one corresponding to the column of the entry. For example, $FTAHub(a)$ is indirectly dominated by $CUGFT$ as there exists a (finite) sequence of outcomes and coalitions such that all coalitions in the sequence prefer the final outcome over the current one:

$$FTAHub(a) \rightarrow_{\{b,c\}} CU(b,c) \rightarrow_{\{a,b,c\}} CUGFT$$

Checking for all possible sequences yields the following indirect dominance matrix:[25]

[25] Appendix A contains the pseudocode for this procedure.

$$
\begin{array}{lc}
 & \begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \end{array} \\
\begin{array}{l}
1\,MFN \\
2\,CU(a,b) \\
3\,CU(b,c) \\
4\,CU(c,a) \\
5\,CUGFT \\
6\,FTA(a,b) \\
7\,FTA(b,c) \\
8\,FTA(c,a) \\
9\,FTAHub(a) \\
10\,FTAHub(b) \\
11\,FTAHub(c) \\
12\,FTAGFT \\
13\,MTA(a,b) \\
14\,MTA(b,c) \\
15\,MTA(c,a) \\
16\,MTAGFT
\end{array}
&
\begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

Note that intuitively any outcome is stable if all deviations from it are deterred. Also, a deviation from the outcome is hindered if a stable outcome might be reached, and some member of the deviating coalition does not prefer it over the initial outcome. In the following procedure, start with the full set and then keep removing unstable elements until the remaining ones are stable.

Take $x \in \{MFN, FTA(i,j), CU(i,j), MTA(i,j)\}$, where $i,j \in N$ with $i \neq j$, and then consider the joint deviation $x \rightarrow_{\{a,b,c\}} FTAGFT$. The $FTAGFT$ regime is not indirectly dominated by any other outcome (see the matrix above) and also $x \prec_{\{a,b,c\}} FTAGFT$ for each of those x. Thus the deviation $x \rightarrow_{\{a,b,c\}} FTAGFT$ cannot be deterred and therefore no such x can be part of the stable set.

Consider $FTAHub(i)$, $i \in N$, and the deviation $FTAHub(i) \rightarrow_{\{j,k\}} FTAGFT$, $j,k \in N \setminus \{i\}$ with $j \neq k$. Using $FTAHub(i) \prec_{\{j,k\}} FTAGFT$ together with the logic from before eliminates $FTAHub(i)$ for each $i \in N$.

Focus on the set of remaining elements $Y = \{CUGFT, FTAGFT, MTAGFT\}$. Start with any element y in Y. If there is a deviation to any element $x \in X \setminus Y$, then there always exists an indirect dominance path (see indirect dominance matrix) $x \ll y'$ coming back to an element $y' \in Y$. In addition, for any $y_1, y_2 \in Y$, $y_1 \neq y_2$, there does not exist a coalition $S \subseteq N$, $S \neq \emptyset$, for which $y_1 \prec_S y_2$. Thus, the set Y satisfies the (internal) stability condition while being maximal, i.e. $Y = LCS$. □

The proposition informs us that in the symmetric case, under the current institutional arrangement of the WTO, (the trinity of) global free trade is the only stable trade policy regime according to our framework.

Yet what would happen in the absence of article XXIV? We now consider an alternative constitutional setup. In that case, countries would not have the option to liberalize trade by forming CUs or FTAs, leaving MTAs as the only possibility, apart from global free trade and MFN. The preference ranking of the representative country a over the reduced set of policy regimes then looks as follows:

$$MFN \prec_a MTA(a,i) \prec_a MTA(i,j) \prec_a MTAGFT$$

Each country achieves its maximal welfare under MTAGFT, i.e., global free trade. Note that FTAGFT and CUGFT are no longer available, nor is FTAHub, the regime that dominated the welfare ranking for the hub country in the presence of article XXIV. It is thus reasonable for us to conjecture the stability of MTAGFT, and the following proposition proves this conjecture:

Proposition 4. *Under symmetry and with the modified institutional arrangement of the WTO (no PTAs), the LCS contains one element: MTAGFT. In other words, global free trade is a unique stable outcome.*

Proof. The indirect dominance matrix is derived as before:

$$\begin{array}{l c}
 & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{l} 1\, MFN \\ 2\, MTA(a,b) \\ 3\, MTA(b,c) \\ 4\, MTA(c,a) \\ 5\, MTAGFT \end{array} &
\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{array}$$

Let us start with the full set of trade agreements again (limited to the setting). If the grand coalition moves from MFN to $MTAGFT$, then the only possibility is to stay there, as any other outcome does not indirectly dominate $MTAGFT$. Moreover, $MFN \prec_{a,b,c} MTAGFT$. Thus, MFN cannot be stable. Furthermore, if the grand coalition moves from any bilateral MTA regime to $GFTMTA$, by the same argument, it is clear that no bilateral MTA can be stable. Finally, any deviation from $MTAGFT$ will come back to itself due to the indirect dominance. Consequentially, the set $Y = \{MTAGFT\}$ is consistent and also the largest one. □

We see that if countries' endowments are symmetric, corresponding to the lower-left corner of the triangle, then the presence of article XXIV does not change the stable outcome: global free trade (or the trinity thereof, if FTAs and CUs are permissible). We defer the consideration of scenarios with only one type of PTA to subsection 3.4.

We now proceed to the analysis of the other two extreme cases, which correspond to points LSS and LSL of the triangle, cf. Figure 4. The acronyms stand for country a being large, country b small (always), and country c either small or large.

We consider first the case of two small and one large country, i.e. point LSS. In this scenario, $e_a = e_{max}$ and $e_b = e_c = e_{min}$. We obtain the following ranking of preferences for country a and for the small countries $i \in \{b, c\}$ and $j \in N \setminus \{a, i\}$):

$$
\begin{array}{ll}
\text{country a:} & CU(i,j), FTA(i,j) \prec_a FTAHub(i) \prec_a MFN, MTA(i,j) \prec_a FTA(a,i) \\
& \prec_a MTA(a,i) \prec_a CU(a,i) \prec_a GFT, FTAHub(a) \\
\text{countries b/c:} & MTA(a,i) \prec_i GFT, FTAHub(a) \prec_i CU(a,i) \prec_i FTA(a,i) \prec_i CU(a,j) \\
& \prec_i MFN, MTA(i,j) \prec_i FTAHub(i) \prec_i FTA(a,j) \prec_i MTA(a,j) \\
& \prec_i FTAHub(j) \prec_i CU(i,j), FTA(i,j)
\end{array}
$$

One immediately notices that small and large countries have different rankings. A large country profoundly dislikes the scenarios where it is an outsider, while the small countries, by contrast, dislike any trade arrangements with the large country. Note that countries do not differentiate between different trade constellations in certain cases.[26] For example, CU(i,j) and FTA(i,j) result in the same welfare for all countries. In this case, under the given pattern of endowments, the optimal tariffs of the small countries for CU and FTA are above the MFN tariff. However, the sub-paragraphs of article XXIV paragraph 5 rule this out, and therefore we cap the tariffs at their MFN level. A similar argument applies to the case of FTAHub(a). Here, the optimal tariffs of the small countries would be negative. Restricting tariffs to be non-negative implies that FTAHub(a) corresponds to GFT, or rather a pseudo-GFT (as external tariffs between spoke countries are zero). Finally, the MTA between the small countries actually coincides with the MFN regime because of identical optimal tariffs for both cases.

Using the above preference rankings to derive the largest consistent set yields the following proposition:

Proposition 5. *With the endowments given by $e_a = e_{max}$ and $e_b = e_c = e_{min}$, and under the current institutional arrangement of the WTO, the stable constellations are the PTAs – CU(b,c) and FTA(b,c) – between the two small countries.*

Even though global free trade is the most desirable regime for the large country, the two small countries do not have an incentive to agree to such a constellation, and the large country cannot enforce it. Consequently, country a ends up with the worst arrangement (from its perspective). So in this scenario, the size advantage of the large country does not translate into a favorable outcome. Moreover, the case at hand demonstrates the relevance of restrictions on PTAs (remember that insiders are not allowed to raise tariffs on outsiders). The constraint renders the small countries indifferent between the two forms of PTAs.

We now turn to the hypothetical scenario without article XXIV paragraph 5. The preference rankings for the large country a and the small countries b and c (indexed by i and j) are as follows:

$$
\begin{array}{ll}
\text{country a:} & MTA(i,j), MFN \prec_a MTA(a,i) \prec_a GFT \\
\text{countries b/c:} & MTA(a,i) \prec_i MTAGFT \prec_i MFN, MTA(i,j) \prec_i MTA(a,j)
\end{array}
$$

[26] Additional details on this can be found in Appendix B.2.

As a result, the best outcome for a small country i is the $MTA(a,j)$ regime, i.e., an MTA between the large country and the other small country, as the PTAs are not available anymore. The next proposition presents the LCS for the reduced set of regimes:

Proposition 6. *With the endowments given by $e_a = e_{max}$ and $e_b = e_c = e_{min}$, and under a modified institutional arrangement of the WTO without article XXIV, the stable constellations are MFN and MTA(b,c).*

In summary, when there are two small and one large country, the GFT regime is unstable under the current and under the hypothetical alternative institutional set-up of the WTO. At best, world trade can be partially liberalized. Furthermore, the small countries benefit when they are allowed to form a PTA instead of an MTA, as the limiting MFN principle can be avoided in that way.

We now turn to the third corner of the triangle, denoted by LSL. In this point, $e_b = e_{min}$ and $e_a = e_c = e_{max}$, i.e., we have one small and two large countries. Let us start with the preference rankings for the small country b and the large countries $i \in \{a, c\}$ and $j \in N \setminus \{b, i\}$:

$$
\begin{array}{ll}
\text{country b:} & GFT \prec_b CU(i,b) \prec_b FTAHub(i) \prec_b FTA(i,b) \prec_b FTAHub(b) \\
 & \prec_b MTA(i,b) \prec_b MFN \prec_b CU(i,j) \prec_b MTA(i,j) \prec_b FTA(i,j) \\
\text{countries a/c:} & CU(j,b) \prec_i FTA(j,b) \prec_i MFN \prec_i MTA(i,b), MTA(j,b) \prec_i FTA(i,j) \\
 & \prec_i MTA(i,j) \prec_i CU(i,j) \prec_i FTAHub(b) \prec_i FTAHub(j) \prec_i GFT \\
 & \prec_i FTAHub(i) \prec_i FTA(i,b) \prec_i CU(i,b)
\end{array}
$$

Under this pattern of endowments, the preference rankings of the countries are considerably different from previous cases. For the small country, the MFN regime generates higher welfare than any other trade agreement of which it is part. As for a large country, being an outsider is at the lower end of its ranking, while being an insider in a PTA with a small country is at the top end.

We determine the LCS under these preference rankings in the following proposition:

Proposition 7. *With the endowments given by $e_b = e_{min}$ and $e_a = e_c = e_{max}$, and under the current institutional arrangement of the WTO, the stable constellation is the CU between the two large countries, that is CU(c,a).*

The small country b manages to block many desirable outcomes for large countries. Country b can unilaterally deviate from any trade agreement with higher welfare than CU(i,j) for the large countries. Thus, despite being a majority, the two large countries cannot impose their will on a single small country. All the large countries can achieve is the best trade agreement that they can reach without the participation of the small country, which is CU(a,c), a customs union among themselves.

A similar story unfolds in the scenario without article XXIV paragraph 5. Under the reduced choice set, the countries' preference rankings are as follows:

country b:	$MTAGFT \prec_b MTA(i,b) \prec_b MFN \prec_b MTA(i,j)$
countries a/c:	$MFN \prec_i MTA(i,b), MTA(j,b) \prec_i MTA(i,j) \prec_i MTAGFT$

As the logic of the respective preference rankings for the countries is similar to before, let us directly present the proposition:

Proposition 8. *With the endowments given by $e_b = e_{min}$ and $e_a = e_c = e_{max}$, and under a modified institutional arrangement of the WTO, the stable constellation is the MTA between the two large countries, that is MTA(c,a).*

Similar to the other asymmetric corner considered before, having one small and two large countries allows for partial but not full liberalization of world trade, irrespective of the constitutional scenario (current vs. modified WTO rules). In terms of overall welfare, the world is better off in the hypothetical scenario without article XXIV, though. Individually, the small country is in a better position in the case of MTA(i,j) relative to CU(i,j), as it exploits the MFN obligation of the large countries. By contrast, the large countries are better off under the PTA. Therefore, while neither of the two institutional arrangements facilitate global free trade; they influence the welfare of the stable set, both globally and its distribution across countries.

3.2. **Edges of the Triangle.** Let us now turn to the edges of the triangle, where the endowments of the three countries vary along a single dimension. Referring back to Figure 4, we denote the southern edge by VSS, where country a varies from small to large, and countries b and c are small, connecting the Sym and LSS corners. The eastern edge is denoted by LSV, where country a is large, b is small, and c varies from small to large, connecting the corners LSS and LSL. The third, northwestern edge, finally, we denote by VSV, where countries a and c vary in unison from small to large, and b is small, connecting the corners Sym and LSL.

Our analyzes of each edge will be summarized by graphs that depict the composition of the stable set as we vary the endowment of one country (or of two in unison in the case of VSV), and we provide accompanying descriptions that explore the underlying forces at play. The exact numerical values for the (sub-)intervals of each edge are provided in Section C.2. To gain an intuitive understanding of the results, we identify specific trade agreements that switch from stable to unstable (or vice versa) at particular endowment tuples, and explore the underlying mechanics to understand the reason for these switches.

3.2.1. *VSS: 1 varying, 2 small countries.* We consider first the scenario where $e_b = e_c = e_{min}$ and $e_a \in [e_{min}, e_{max}]$ (the edge connecting corners Sym and LSS). Alongside this edge, certain trade policy arrangements never appear in the LCS, in particular the MFN, FTAHub(a), and MTA(b,c) regimes. The stable sets along this edge are presented in Figure 5.[27]

[27]Where the dot marks a single point, and the height of different lines in the diagram has no meaning except for keeping the elements of the LCS distinguishable as we move along the edge horizontally.

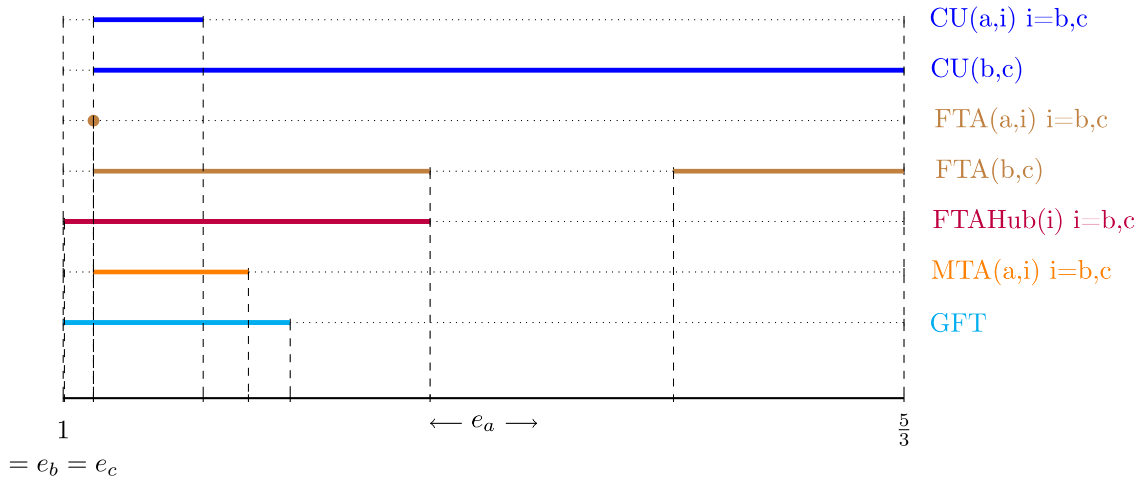

FIGURE 5. Characterisation of varying, small, and small country

In the left corner (Sym), where e_a is close to one and the size distribution close to symmetry, the GFT regime (that is, the trinity of FTAGFT, MTAGFT, and CUGFT) is the only element of the LCS. As e_a increases slightly, both FTAHub(b) and FTAHub(c) emerge as stable outcomes. Along the whole interval, a number of different PTA and MTA constellations, either between the growing country a and one of the smaller countries (b/c) or between the two smaller countries b and c appear. Towards the upper end, where e_a tends to 5/3 and we reach the LSS corner, only the PTAs between the small countries b and c are still stable.

First, the spike in the number of stable constellations close to symmetry actually follows from a change in the preferences of the varying country with respect to CU(b,c) vs. the trinity of GFT regimes — it starts preferring the former over the latter as it grows larger. Next, FTA(a,b) and FTA(c,a) become unstable because the small countries start to like the MFN regime more than the GFT trinity (or rather, these arrangements are only stable for as long as this is not the case). When country a becomes sufficiently large, country b prefers FTAHub(c) over CU(a,b), and c prefers FTAHub(b) over CU(c,a). As a consequence, both of these CUs drop out of the LCS. Similarly, when countries b and c start preferring FTAHub(c) and FTAHub(b) over GFT, the latter stops being stable. A similar argument also applies to the MTA regimes. When the size of country a increases even further, both country b and c favor CU(b,c) over the respective hub constellations, which results in FTA(b,c), as FTAHub(b) and FTAHub(c) become unstable. When the endowment of country a tends towards its maximum, the small countries are constrained by the MFN-tariffs and do not differentiate between CU(b,c) and FTA(b,c) anymore, which renders FTA(b,c) stable again.

We now turn to the hypothetical scenario without Article XXIV Paragraph 5. The interval over which the GFT (MTAGFT) regime is stable increases significantly. Moreover, over two-thirds of the interval, the GFT regime is the unique element in the LCS. Additionally, all possible combinations of MTA appear at some point (mostly close to the symmetry). Figure 6 demonstrates these results.[28]

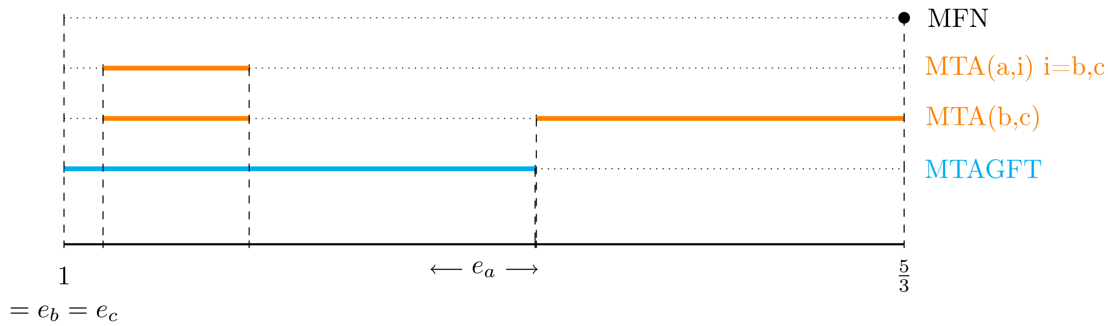

FIGURE 6. Characterisation of varying, small, and small country

Close to symmetry, MTAGFT is the only element in the stable set. As soon as the small countries start to prefer MTAs with country a over MTAGFT, all three MTAs appear in the LCS. As the size of country a increases, the MTAs drop out from the LCS because the small countries rank the MTA with the large country as the worst trade agreements (switching last place with the MFN regime), which actually also influences the stability of the MTA among themselves. Furthermore, the GFT regime becomes unstable when the small countries start to prefer their joint MTA over MTAGFT.

To conclude, the effect of the PTAs on the stability of the GFT regime is significant: the abolishment of Article XXIV Paragraph 5 would facilitate the formation of GFT as long as there are two small countries and the third country is not substantially larger.

3.2.2. *VSV: 2 varying, 1 small country.* Now, let us turn to the case where $e_b = e_{min}$ and $e_a = e_c \in [e_{min}, e_{max}]$ (edge VSV). Along this edge, depending on the size of the larger countries (a/c), any trade agreement can be part of the stable set. The exact composition of the LCS is indicated by Figure 7.

Close to the symmetric corner (Sym), the trinity of GFT regimes is stable and remains the unique element(s) of the stable set for longer (compared to the previous case, VSS). In addition, a collection of different trade agreements is stable relatively close to symmetry. Towards the other side, as $e_a = e_c$ tends to 5/3 (corner LSL), the CU between the two (varying) countries a and c is the unique stable outcome. Furthermore, the MFN regime is also an element of the LCS in two small (one tiny), separated regions.

Again, the peak in stability near symmetry stems from a shift in the preferences of the varying countries (a and c) with respect to the CUs. At the point of the shift, both these countries start to prefer a respective CU with the small country over the trinity of GFT. The occurrence of the MFN regime follows from a preference of the small country (b) of MFN over the trinity of GFT (the first region), and then over FTAHub(a) and FTAHub(c)

[28]As before, in addition to the above-mentioned trade agreements, the graphic also has MFN as a single point, see the dot, at $e_a = 5/3$ (point LSS).

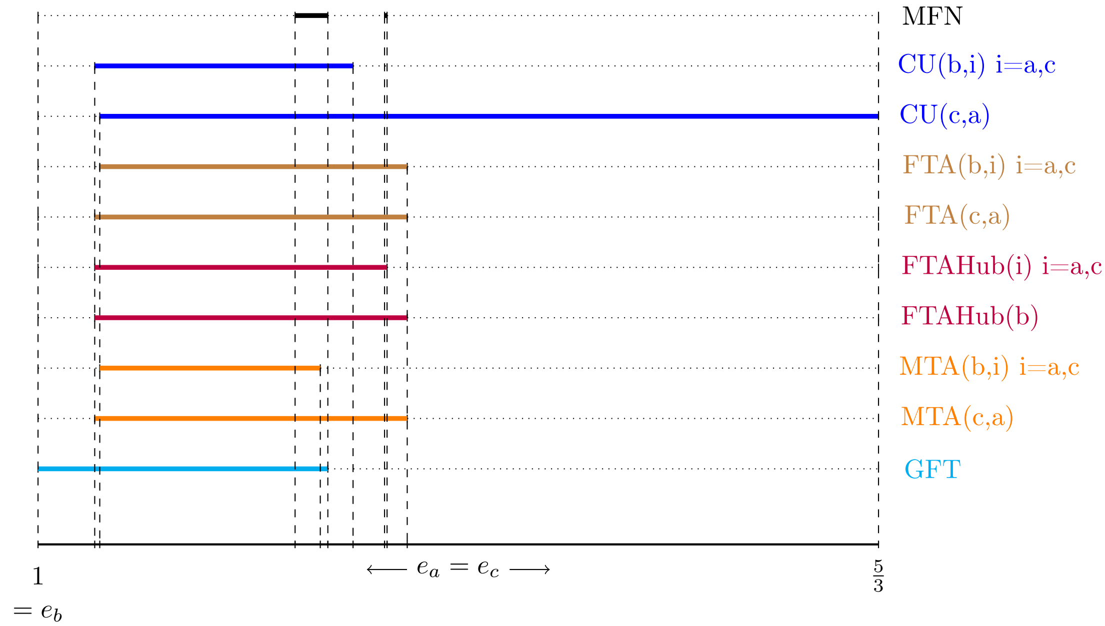

FIGURE 7. Characterisation of varying, small, and varying country

respectively (the second, tiny region). As countries a and c grow larger, first MTA(a,b) and MTA(b,c) drop out from the LCS as they rank lowest according to the preferences of country b, and then MTA(a,c) follows. The trinity of GFT becomes unstable once the small country b starts to prefer CU(c,a). Next, CU(a,b) and CU(b,c) follow in dropping out of the LCS as the small country starts to prefer the MFN regime over a CU with any of the larger two countries. As soon as CU(c,a) becomes preferred by country b over any FTA where b participates, or any hub structure with a large country as the hub, all aforementioned constellations drop out from the LCS.

We now switch off Article XXIV Paragraph 5 and consider the alternative scenario without PTAs. In contrast to the previous case (edge VSS), the interval where the GFT regime (MTAGFT) is part of the stable set decreases. However, the reduction in range is slight. A similar observation holds for the range close to symmetry, where the GFT regime is the unique stable outcome. The exact composition of the LCS is shown in Figure 8.

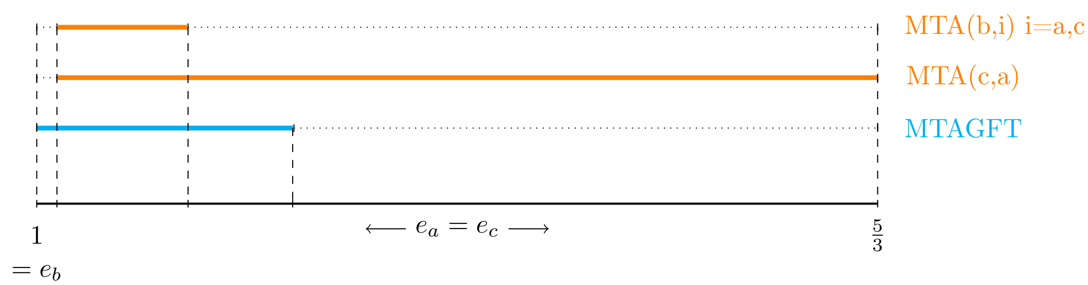

FIGURE 8. Characterisation of varying, small, and varying country

The main driving force behind the changes in the set of stable outcomes are alterations in the preferences of the small country over the interval. More precisely, it is essential where exactly the small country places the

MFN regime in its ranking compared to the other trade agreements. As soon as country b prefers MFN over another constellation, the latter drops out from the LCS. Beyond a certain point, the MTA between the large countries a and c remains the only stable outcome.

Similar to the previous edge (VSS), the case at hand renders a precise prediction of stable outcomes difficult, as it is hard to rule out regimes, especially around symmetry where, as before, almost all constellations are stable. The effect of the abolition of Article XXIV Paragraph 5 affects the stability of the GFT regime in the opposite direction when compared to the previous case (VSS).

3.2.3. *LSV: 1 large, 1 small, 1 varying.* Finally, we turn attention to the third edge of the triangle where we find a unique composition of LCS under the hypothetical scenario without PTAs. As before, though, we start by considering the status-quo institutional setup with PTAs. Along LSV, $e_a = e_{max}$, $e_b = e_{min}$, and $e_c \in (e_{min}, e_{max})$. For these endowment constellations, several trade agreements can be ultimately ruled out (with respect to the LCS). The MFN and (trinity of) GFT regimes, for example, are never part of the stable set. Additionally, none of the PTAs between the small and the large country appear as a stable outcome. The same holds for the FTA-hub structures where either the small or the large country is the hub node. As for the actual composition of the LCS along this edge, see Figure 9 for a graphical representation.

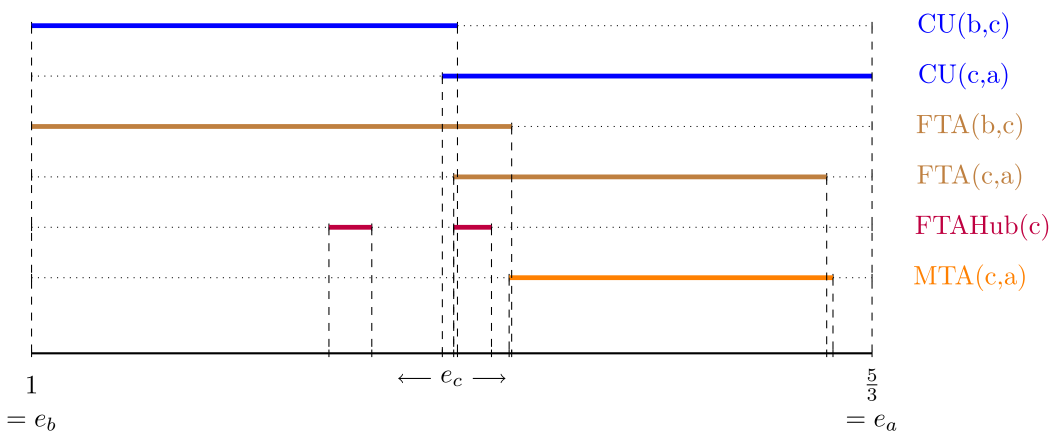

FIGURE 9. Characterisation of large, small, and varying country

The general observation is that when the varying country c is close in size to the small country b, then the PTAs between these smaller countries appear as elements in the stable set. As the varying country c becomes larger, the trade policy constellations between the larger countries a and c replaces these. Additionally, there are two small, separate regions in the middle of the interval where FTAHub(c) is stable.

Start with the PTAs between countries b and c, the small and the varying one. Interestingly, the only factor driving the stability of these policy regimes are the preferences of country b (with fixed minimal endowments). Once the MFN regime becomes more desirable than CU(b,c) for country b, the constellation CU(b,c) drops from the stable set. Now, an identical line of reasoning holds for the case of FTA(b,c). Thus, for both constellations, it only takes a single change in the preference ranking of country b to alter the stable set.

The PTAs and MTAs between countries a and c start to appear in the LCS when country c is becoming relatively large and closer to country a in size. At first, both countries actually prefer to form a CU with the small country b, that is when country c is relatively small (and CU(b,c) actually is an element in the stable set). However, once it becomes preferable for country b to be the outsider instead of the insider in a CU, CU(c,a) emerges as a stable outcome (even though CU(b,c) still remains stable). Moreover, as soon as country c prefers FTA(c,a) respectively MTA(c,a) over the MFN regime, each of them becomes part of the LCS as well. For the interval where all PTAs and MTAs between country a and c are stable, both countries have fixed preference relations over these outcomes:

country a:	$FTA(c,a) \prec_a CU(c,a) \prec_a MTA(c,a)$
country c:	$MTA(c,a) \prec_c FTA(c,a) \prec_c CU(c,a)$

However, as soon as country c also prefers MTA(c,a) over FTA(c,a), the joint FTA drops out of the LCS. Similarly, as soon as country a prefers CU(c,a) over MTA(c,a), this also applies to the joint MTA – CU(c,a) arises as the only stable outcome.

FTAHub(c) is stable in the two small, separate regions in (or near) the middle of the interval. In the first region, the stability is driven by the fact that country b starts to value FTAHub(c) more than FTA(b,c) and hence agrees with country a's preferences in this respect. Once the preferences of country b over these outcomes are reversed, FTAHub(c) drops out of LCS again. In the second region, the stability of the same hub structure is mainly determined by the change in the preferences of country c. Now, once it starts to value FTA(c,a) over the MFN regime, which also puts FTA(c,a) in the LCS; both FTAs with c as a partner are stable and by consequence, the corresponding hub structure is stable as well. As the free-riding incentives of the country b increase (valuing the MFN regime more than FTAHub(c)), this hub structure is no longer part of the stable set.

The alternative institutional scenario without Article XXIV does not promote the appearance of GFT as part of the stable set. GFTMTA, but also MTA(a,b) and MTA(b,c) never emerge as stable outcomes. Varying the size of country c generates either the MFN regime or MTA(c,a) as the stable elements. Figure 10 presents these findings.[29]

Over the whole interval, country b does not have any incentive to form an MTA with either of the other countries. This is one reason why the MFN regime is stable for part of the interval. The other reason is that country c prefers not to have a trade agreement with country a as long as its size is not large enough. Once

[29] In addition to the elements above, it also depicts MTA(b,c) as a single point, see the dot, but this appears only for the sake of completeness because that point corresponds to the corner LSS discussed earlier.

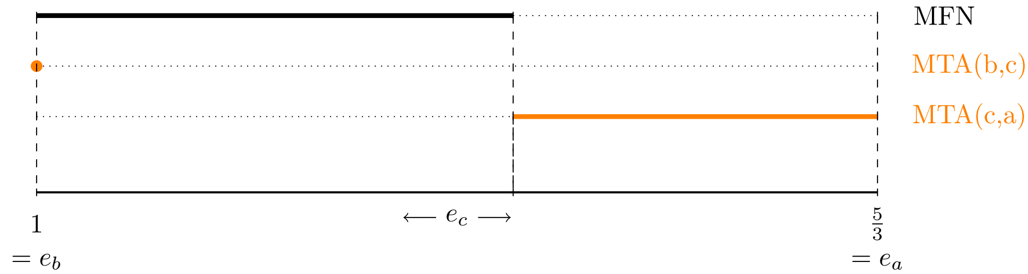

FIGURE 10. Characterisation of large, small, and varying country

country c is sufficiently large, MTA(c,a) presents a better option than the MFN regime. As a consequence, MTA(c,a) replaces the MFN regime as the single element of the stable set.[30]

In conclusion, for the endowment constellations along this edge, the GFT regime never appears as part of the stable set, independent of the scenario (with/without PTAs). However, the constitutional setup determines whether partial trade liberalization may occur or not. The possibility of forming PTAs reduces the incentive of the small(est) country to free ride. Otherwise, without PTAs, the MFN regime is the unique stable outcome when there is one small, one large, and one rather small country, and it is the worst outcome in terms of overall welfare. Prohibiting PTAs has this important and negative implication only along this edge.

3.3. **The interior of the triangle.** We now focus on the interior of the triangle in Figure 4. That is, we now consider the full parameter space of endowment allocations. For expositional clarity, we group both CUs and FTAs (incl. FTA-Hubs) together under the label of PTA in figures 11-13.[31] For the same reason, we also suppress the exact member countries of specific trade agreements, i.e., who is an insider and who stays out.

As before, we start by considering the existing institutional set-up, where PTAs are options countries can choose. Figure 11 depicts regions of the parameter space and their stable sets (or rather a simplified view of these). In a small region close to symmetry, labeled as region 1, the (trinity of) GFT regime(s) is the unique stable element. In an adjacent and a slightly more distant area, both labeled as region 2, PTAs become stable as well. The region connecting those two, which we denote as region 3, adds MTAs as another element of the stable set. In a tiny area along the diagonal, indicated as region 4, no form of trade agreement can be excluded from the stable set.[32] Towards the asymmetric corners LSS and LSL, we find two regions, both labeled as 5, where PTAs are the only stable trade policy constellation. In between those two lies region 6, where MTAs are also found to be stable. Finally, in another tiny area along the diagonal, indicated as region 7, MFN enters the stable set as well. In general, with a certain degree of asymmetry between countries' endowments, at most, partial trade liberalization can be expected, as (the trinity of) GFT is an element of the stable set only in regions 1 through 4, which are all characterized by relatively symmetric endowment allocations.

[30]Note that MTA(b,c) is stable in the LSS corner, that is, at the lower bound of the interval.
[31]However, graphs of this analysis that distinguish between the two types of PTAs can be found in Appendix **??**.
[32]Note that this does not imply that every CU, every FTA or FTA-Hub combination of countries is stable.

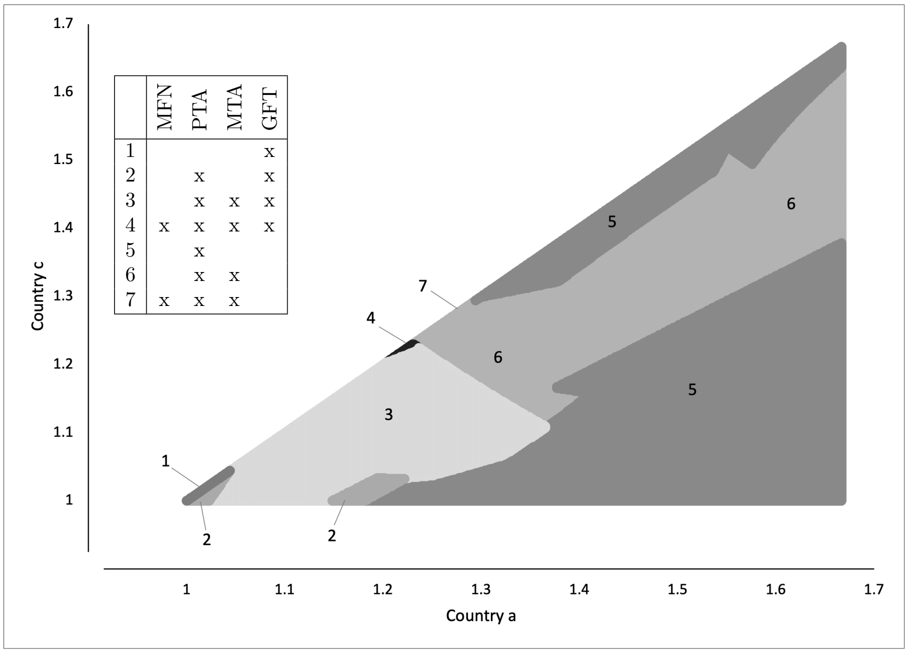

Figure 11. Aggregate View of Overall Stability with PTAs

Next, we consider the alternative institutional setup, where PTAs are not allowed. Figure 12 depicts regions of the parameter space and their corresponding stable sets. In a small area near symmetry as well as in a sizeable area away from it, both denoted as region 1, GFT (that is, MTAGFT here) is again the unique stable element. Adjacent to these are two areas, both marked as region 2, where MTAs become stable as well. Moving towards the asymmetric corners, two regions denoted by 3 feature MTAs as the only stable element. In between those, in region 4, only the MFN is the unique element of the stable set.

The comparison of the two graphs allows us to deduce two compelling statements. The first noteworthy result is the extent of MFN under each scenario. In the alternative institutional setup without PTAs, the area where MFN is an element of the stable set (in fact, the unique element in this scenario) is substantially larger than under the status-quo institutional rules. Note that this effect is at work sufficiently away from symmetry towards the corner LSS. Under (significant) asymmetry, then, it seems that PTAs allow countries to move towards their international efficiency frontier by forming such agreements (cf. Bagwell et al. (2016)). The reason for this is that PTAs allow countries to discriminate against the outsider, whereas an MTA does not allow for this.

The second interesting result is the difference in the extent of stability of GFT across the two regulatory scenarios. First, recall that once the degree of asymmetry exceeds a certain threshold, none of the GFT regimes remain in the stable set, irrespective of the institutional set-up. Around symmetry, by contrast, the opposite holds in that the GFT regimes are always stable (under both scenarios). In between, the effect of PTAs on the stability of GFT depends on the structure of asymmetry. Figure 13 depicts the different areas of stability of the GFT regimes, depending on the regulatory scenario — with or without PTAs. Note that region 1, close

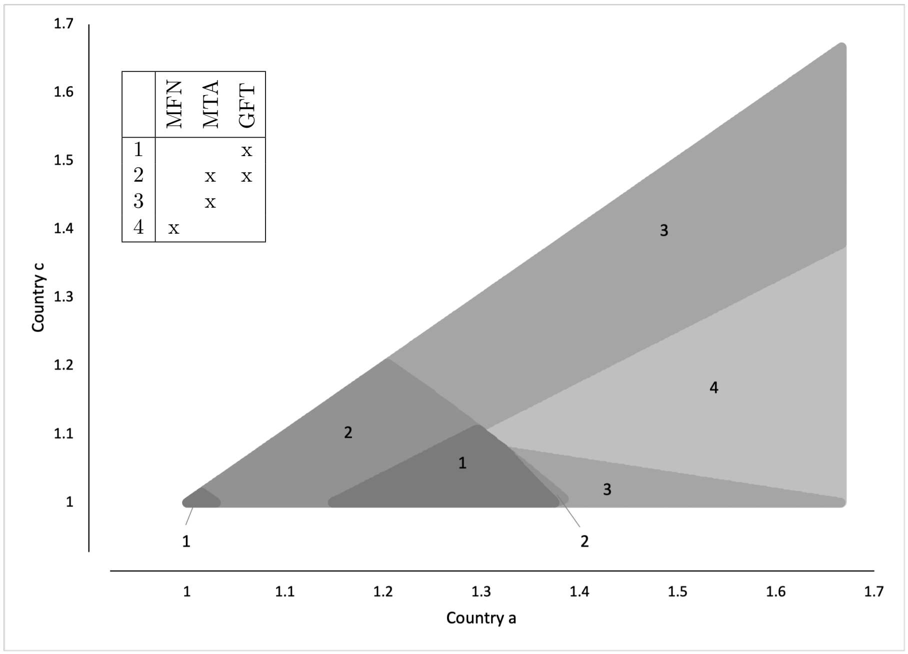

Figure 12. Simplified Overall Stability without PTAs

to symmetry, indicates the stability of GFT under both regulatory scenarios, as described above. Away from symmetry, in case two countries are relatively larger (but not too large), the abolishment of PTAs results in reducing the area where GFT is stable, corresponding to region 2 in the diagram. In this region, PTAs act as 'building blocks' on the road to GFT, as they generate sufficient costs for the small country in the case of leaving the a GFT arrangement (either CUGFT or FTAGFT here). So punishment in the form of a discriminatory external tariff of a PTA between the relatively large countries overcomes the free-riding incentive of the small country. By contrast, if two countries are relatively smaller (but not too small), the same regulatory difference yields the exact opposite effect, see region 3. Here, PTAs act as 'stumbling blocks'. When two countries are relatively smaller, a PTA arrangement between them increases their incentive to exclude the single large country. If they form an MTA (when PTAs are not allowed), they would not have such an exclusion incentive, as the non-discriminatory principle applies. The comparison of the two different forms of asymmetry thus yield an important insight: whether PTAs turn out to be 'building blocks' or 'stumbling blocks' for intermediate levels of asymmetry depends on the relative size of two countries versus country three.

We now go one step further and ask: whenever PTAs are conducive to global free trade, which form of PTA is responsible for the result? In other words, we want to find out which form of PTA, a CU or an FTA, contributes to the stability of GFT when PTAs are available (regions 1 and 2 of Figure 13). The subsequent figure is meant to shed light on this question. From Figure 14, one can see that the stability for GFT (due to the option of forming PTAs) in region 2 stems from the possibility to form a CU among the relatively larger countries. For endowment combinations with somewhat larger countries, as in region 2 of Figure 14, an FTA would not

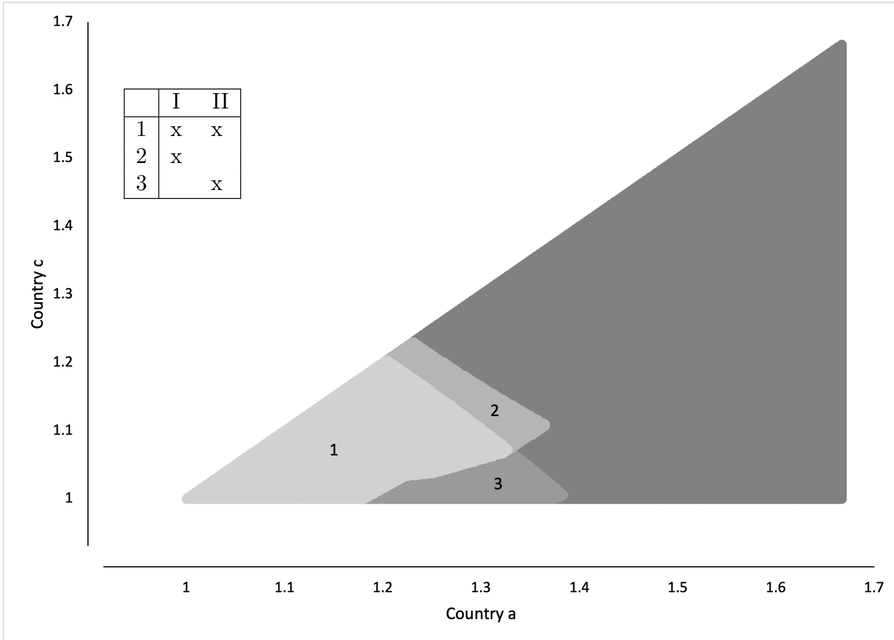

Figure 13. Stability of GFT regimes with (I) and without (II) PTAs

generate enough incentive for a small country to be a member of a GFT arrangement. Under the FTA, member countries impose externalities on each other, resulting in lower tariffs vis-á-vis an outsider (the small country). While under a CU, the larger countries coordinate their decisions, and the resulting tariffs against the small country are higher. The higher tariffs incentivize the small country to be a member of CUGFT. Contrary to region 2, in region 1 of Figure 14, both forms of PTAs guarantee that a GFT regime is stable. Note that here we have always analyzed an alternative scenario where both forms of PTAs are available. In the following (and final) section (of analysis), we consider the CU vs. FTA question in scenarios where we switch off one of them only, one at a time.

Summing up our results so far, if the world is characterized by an intermediate level of asymmetry, with two out of three countries close to identical in size while relatively smaller than the third country, the area where GFT is stable increases when prohibiting PTAs. However, when the two similar countries are larger (not too large), the availability of PTAs (in particular, a CU between the two larger countries), is conducive to the GFT regime's stability. If the world is even further away from symmetry, full trade liberalization is not attainable at all, and an area where the MFN regime is stable appears under the scenario without PTAs.

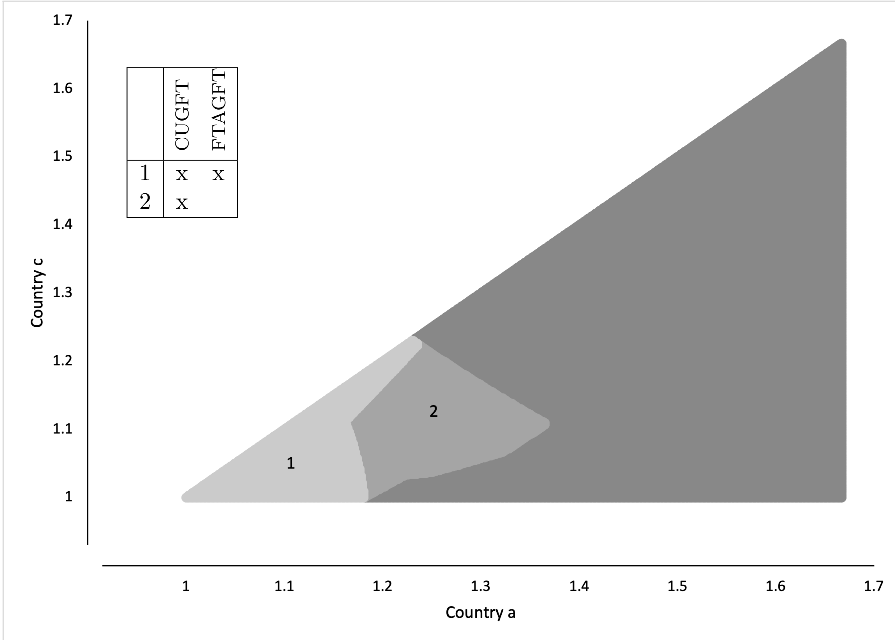

FIGURE 14. Stable area for GFT reached via CUs/FTAs

3.4. **Hypothetical scenarios with only FTAs or CUs.** Figure 13 showed that the availability of PTAs can increase the stability of GFT (region 2). Figure 14 then indicated that this finding is primarily due to customs unions. We obtained this insight in a scenario where both forms of PTAs are available to decision-makers. Taking the analysis one step further to gain a better understanding, we now extend the space of constitutional setups by considering two additional hypothetical scenarios: In the first scenario, FTAs are ruled out, so countries can form MTAs and CUs only, while in the second scenario, CUs are ruled out, so countries can create MTAs and FTAs only. So we now have four scenarios: status-quo, no-FTA, no-CU, and no-PTA

Focusing on no-FTA and no-CU, we demonstrate the respective results in Figures 15 and 16. Figure 15 shows the results for the hypothetical scenario where FTAs are not available. In Figure 16, we represent the regions of the parameter space and their corresponding stable sets for the hypothetical scenario when we rule out CUs.

A few points are worth noting. First, across both scenarios, there is no endowment configuration for which MFN is a stable outcome. Second, even though the area where GFT is uniquely stable is more extensive in a scenario with only CUs compared to the one with only FTAs (region 1 in Figure 15 vs. region 1 in Figure 16, the latter consisting only of the point of symmetry), the total area where the stability of GFT cannot be ruled out is larger without CUs (regions 1 and 2 in Figure 15 vs. regions 1, 2 and 3 in Figure 16). Once we move further away from the point of symmetry, the scenarios considered here do not generate qualitatively different results compared to the scenario with both types of PTA available (cf. Figure 11): towards the asymmetric right-hand-side of the triangle, only partial liberalization is attainable.

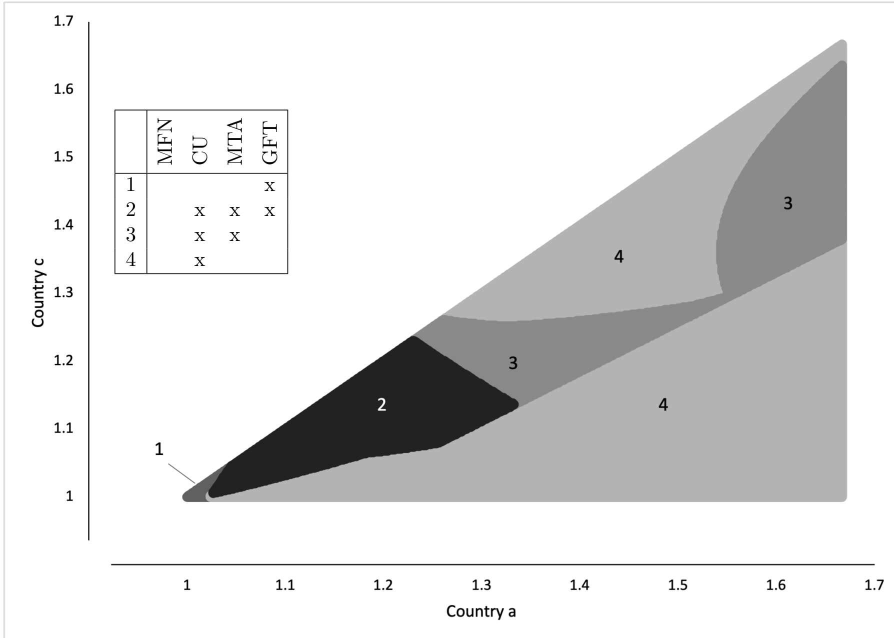

Figure 15. Aggregate View of Overall Stability without FTAs

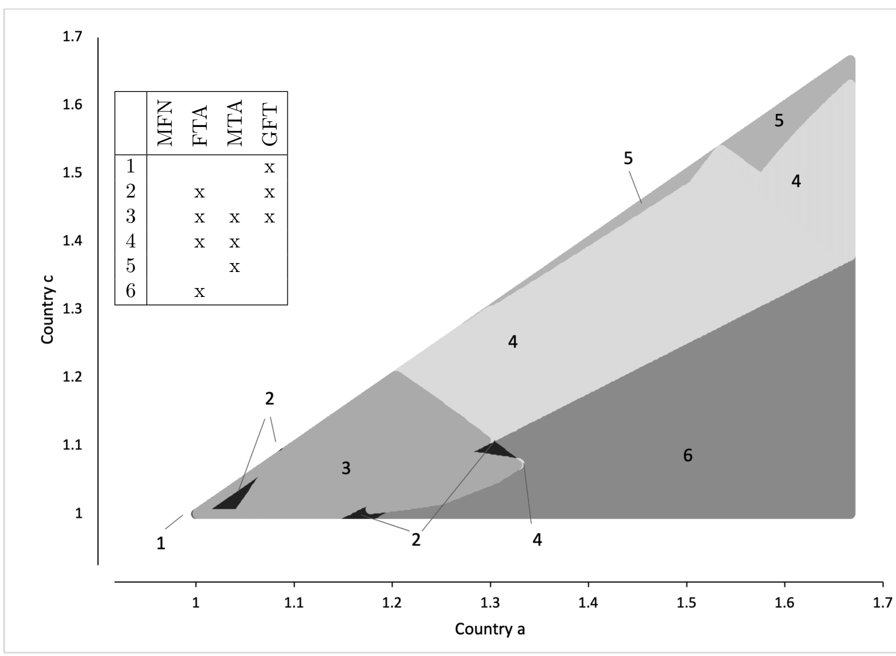

Figure 16. Aggregate View of Overall Stability without CUs

We now turn to the stability of GFT for the four scenarios under consideration: a scenario with PTAs (the existing status-quo institutional set-up), the alternative scenario without PTAs (only MTAs available), and in addition, a scenario with only CUs (no FTAs available) as well as a fourth scenario with only FTAs (no CUs available). Figure 17 demonstrates the regions where GFT is a stable outcome under each of these scenarios.

The fundamental trade-off between the stability of GFT regimes with and without PTAs (regions 2 and 3 in Figure 13) is also present in a scenario with only CUs (no FTAs allowed). In this scenario, if we turn off the possibility of forming CUs, there is an increase in the area where GFT is stable (regions 2, 3, and 4 in Figure 17), but we lose region 5, where GFT is no longer stable. By contrast, in a scenario with only FTAs (no CUs allowed), doing away with FTAs increases the area where GFT is stable (region 4 in Figure 17). So without free trade agreements, global free trade remains stable for a larger constellation of endowments, suggesting that free trade agreements serve as a 'stumbling bloc'. A further important consideration is the comparison between these two scenarios in relation to the status quo, where both forms of PTA are available. On the one hand, when we switch off the possibility of forming CUs, even though there is a slight gain in the area where GFT is stable (region 3 in Figure 17), we lose regions 5 and 6, where the GFT regime is no longer stable. If we turn off FTAs, on the other hand, we do not gain any area where GFT is stable, but instead lose stability in areas 2 and 6 of Figure 17. Consequently, the joint availability of FTAs and CUs may be the most effective policy response to address the smallest country's free riding incentive or the exclusion incentive of the smaller countries.

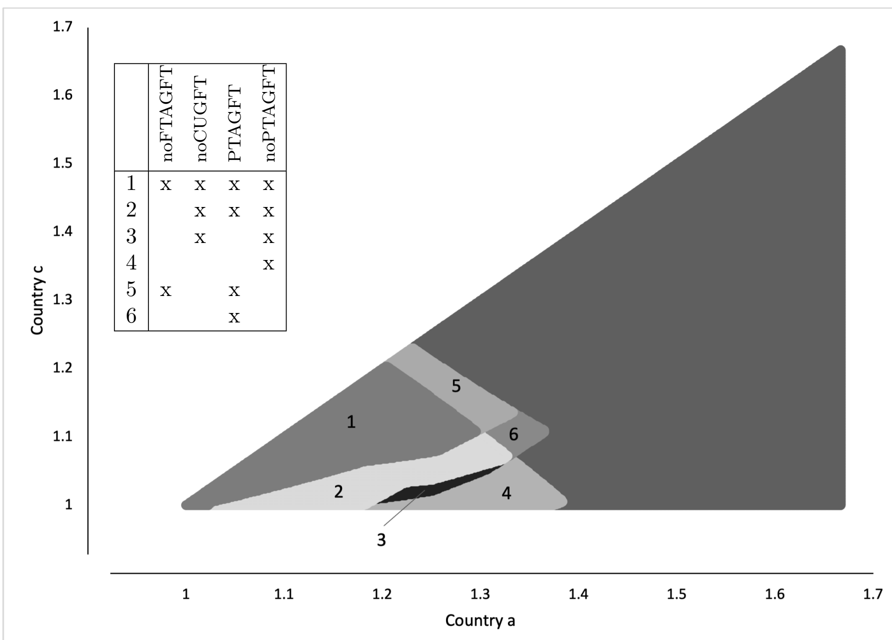

FIGURE 17. Different Scenarios and Stability of Global Free Trade Arrangements

We want to explore further the interaction between both types of PTA in reaching global free trade. In the previous subsection, we discussed region 2 of Figure 13 (the upper north-east part of which is made up of regions 5 and 6 of Figure 17), and concluded that the availability of CUs is essential for extending the stability of the GFT regime to region 2. Now we can present a more nuanced view: In region 5 of Figure 17, the stability of GFT is entirely dependent on the availability of CUs. In region 6 of Figure 17, however, the availability of FTAs is required in addition for CUs to play a similar role. More precisely, comparing constellations of country endowments that fall into region 5 as opposed to those in region 6 in the scenario where only CUs are available (no FTAs), an interesting pattern emerges. The middle-sized country, c, begins to favor the MFN regime over the MTA between itself and the largest country a. For the smallest country b, the MTA between a and c is the most beneficial policy regime, providing a strong incentive for free riding. The change in ranking of MFN vs. MTA (between countries a and c) from the perspective of the middle-sized country c as we switch from area 5 to 6 results in the CU between countries b and c becoming indirectly undominated, and it is the only regime that is indirectly undominated. As a consequence, an exclusion incentive is triggered, resulting in the CU between countries b and c becoming the only stable outcome, leaving the largest country as an outsider and effectively preventing GFT from being stable. If now we allow FTAs, the Customs Union between countries b and c loses its status of indirectly undominated. It becomes indirectly dominated by various policy regimes involving FTAs, which, in turn, are themselves indirectly dominated by GFT. Consequently, global free trade is stable when both forms of Preferential Trade Agreements are available.

In summary, within region 6 depicted in Figure 17, the stability of global free trade is guaranteed by the strategic complementarity between CUs and FTAs. In addition to its economic substance, this result also emphasizes the importance of allowing the full range of possible trade policy arrangements when studying their stability, as the analysis could never account for such interactions if FTAs or CUs were excluded ex ante.

4. Comparison with literature

Before concluding this paper, let us discuss the importance of farsightedness in determining the stable set of trade policy arrangements. Due to differences in the set of trade policy regimes that are considered by the related literature, it is not always possible to compare the results in detail. In specific cases, however, we are able to compare the composition of the stable sets in our paper with the results found by Saggi, Woodland, and Yildiz (2013), who apply Coalition Proof Nash Equilibrium (CPNE) to determine the stability of trade policy constellations in their paper. Their "multilateralism game" is essentially identical to our scenario of a modified institutional setup that excludes PTAs. Therefore, Figures 2 and 5 in Saggi, Woodland, and Yildiz (2013) can be directly compared to Figures 8 and 6 in this paper.[33]

As a starting point, we shall discuss the edge VSS, i.e. the case of two small countries and one (larger) country of varying size (country a being the country varying in size, while countries b and c are the small countries). Close to the boundaries of the interval, i.e. near the corners labelled Sym and LSS, the results of both studies coincide, despite the difference in solution concept. That is, when the size of the larger country a falls within the intervals [1, 1.030] or [1.375, 1.667], both the current study as well as Saggi, Woodland, and Yildiz (2013) predict as the unique outcome MTAGFT (for $e_a \in [1, 1.030]$) or an MTA between the small countries b and c (for $e_a \in [1.375, 1.667]$).

[33] Be aware that their figures are not drawn to exact scale, i.e. the length of the segments is conceptual.

As soon as the size of the varying country exceeds 1.030, the unilateral deviation incentive of a small country (say b) from MTAGFT to $MTA(c,a)$ is decisive for the stability of MTAGFT under the concept of CPNE. Based on Saggi, Woodland, and Yildiz (2013), a small country's incentive to deviate from global free trade to $MTA(c,a)$ prevents the MTAGFT regime from remaining stable. We demonstrate in our paper that under the LCS solution concept, such deviation incentive will not be sufficient to prevent MTAGFT from remaining stable, because the $MTA(c,a)$ regime and all other regimes remain indirectly dominated by MTAGFT. In particular, we cannot preclude global free trade from being stable due to the following indirect dominance paths:

$$MTA(c,a) \rightarrow_{\{c\}} MFN \rightarrow_{\{a,b,c\}} MTAGFT$$

or

$$MTA(c,a) \rightarrow_{\{c\}} MFN \rightarrow_{\{b,c\}} MTA(b,c) \rightarrow_{\{a,b,c\}} MTAGFT$$

If one of the small countries (say, country b) diverges from global free trade, one might expect that the other small country (country c) would also diverge from $MTA(c,a)$ to MFN, even though country c would be worse off under MFN. Yet since all other countries would also be worse off under the MFN regime, decision makers in country c may believe that any one of the indirect dominance paths mentioned above is feasible. In light of this, one should not rule out the possibility of MTAGFT remaining stable. Furthermore, MTAGFT is indirectly (even directly) dominated by an MTAs between the varying (large) and either one of the small countries, while all possible constellations of 2-country MTAs indirectly dominate one another. As long as such indirect dominance relations are established in the interval [1.031, 1.147] for varying country size of country a, the LCS will contain MTAGFT as well as all three 2-country MTAs — in contrast to CPNE, which excludes global free trade.

Next, for the interval $e_a \in [1.148, 1.374]$, Saggi, Woodland, and Yildiz (2013) conclude that the MTA between the two small countries is the only stable outcome. By contrast, under the concept of the LCS, global free trade is the only stable outcome. As the results are qualitatively different, we discuss the reasons for this difference in greater detail. We begin by presenting the ranking of policy regimes for all countries within the given interval:

country a:	$MFN \prec_a MTA(b,c) \prec_a MTA(a,i) \prec_a MTAGFT$
country b,c:	$MTA(a,i) \prec_i MFN \prec_i MTA(i,j) \prec_i MTAGFT \prec_i MTA(a,j)$

In addition, we illustrate the indirect dominance relationships.[34]

$$\begin{array}{l c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{l} 1\,MFN \\ 2\,MTA(a,b) \\ 3\,MTA(b,c) \\ 4\,MTA(c,a) \\ 5\,MTAGFT \end{array} & \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

[34] Over the interval in question, both the preference rankings as well as the dominance relations remain unchanged.

When one considers the preference rankings of the countries over trade policy regimes, it is straightforward to see how Saggi, Woodland, and Yildiz (2013) arrive at their conclusion. Free trade is ruled out under CPNE since any of the two small countries b/c would have an incentive to deviate, and such deviation would be self-enforcing. This argument is also valid for $MTA(a,b)$ or $MTA(a,c)$, as small countries generally prefer the MFN regime over an MTA with the large country. At the same time, the MFN regime itself is not stable because the small countries have a joint incentive to deviate from MFN and form an MTA ($MTA(b,c)$). This leaves them to consider $MTA(b,c)$. The joint deviation of a large and a small country does not occur, since the small country prefers the other small country over the larger country as a partner. Moreover, none of the small countries have any unilateral incentive to deviate from the $MTA(b,c)$ regime. This leaves only the joint deviation of all three countries to free trade. The first argument above, however, excludes this deviation. Hence, they find $MTA(b,c)$ to be uniquely stable.

Is the CPNE-induced prediction convincing? In our opinion, it is not. Consider the case in which one of the small countries (say b) deviates from free trade (such deviation is self-enforcing under CPNE) to $MTA(c,a)$. Another small country (c in this case) strictly prefers the MFN regime over the $MTA(c,a)$ regime. But will country c move to the MFN regime? Country c has a short-sighted incentive to move to the MFN regime, as she prefers it over $MTA(c,a)$. Importantly, country c can also move to MFN from a far-sighted perspective, as MFN cannot be a stable outcome. According to the indirect dominance relation given in the matrix above, there are two possible paths that might follow from it:

$$MFN \rightarrow_{\{b,c\}} MTA(b,c) \rightarrow_{\{a,b,c\}} MTAGFT$$

or

$$MFN \rightarrow_{\{a,b,c\}} MTAGFT$$

Both welfare improving paths lead to global free trade. Therefore, no small country will deviate initially from MTAGFT. In other words, all trade policy constellations are indirectly dominated by the MTA between the two small countries and MTAGFT, where the former regime (MTA between the small countries) is indirectly dominated by MTAGFT. If any of the small countries were to deviate from global free trade to $MTA(a,b)$ or $MTA(c,a)$, we should expect further deviations. While indirect dominance paths can lead to MTAGFT or $MTA(b,c)$, the MTA between the small countries is in turn indirectly dominated by global free trade. Consequently, deviations from MTAGFT are deterred, and global free trade is the only stable outcome under our solution concept of LCS. In summary, under CPNE, country b is encouraged to deviate from free trade to $MTA(c,a)$ since no further sub-coalition's deviation exists. Using the concept of LCS, by contrast, country b must take into account what other countries can do further down the line, and those possible moves by other countries deter the initial deviation of country c.

Let us now consider the case of one small (country b) and two larger countries (a and c) of varying sizes. As for the edge considered previously, both approaches predict the same stable sets near the interval's endpoints (corners Sym and LSL). As long as e_a and e_b belong to the interval [1, 1.015], global free trade is a unique stable outcome under the CPNE and LCS solution concepts. In addition, both our paper and that of Saggi, Woodland, and Yildiz find that an MTA between large countries is the only stable outcome in the interval [1.203, 1.667]. There is, however, a substantial difference in stable sets in the interior along this edge. CPNE suggests that as soon as a small country has the incentive to deviate unilaterally from MTAGFT (which happens when the size of large countries is larger than 1.015), this regime ceases to be stable. Despite the shortsighted motive for departing from global free trade, we find that global free trade remains part of the stable set ($e_a, e_b \in [1.015, 1.203]$). As long as MTAGFT indirectly dominates all other trade policy regimes, we cannot exclude it from the stable set. MTAGFT drops out of the LCS once it ceases to indirectly dominate the MFN constellation.

Having compared our results to those of Saggi, Woodland, and Yildiz (2013), for those specific cases where a direct comparison is possible, we observe that the differences between limited (CPNE) and unlimited farsightedness (LCS) boil down to the following: 1. the length of the range over which MTAGFT is stable, and 2. differences between the predicted stable outcomes. One can speculate that for specific subsets of the endowment distribution, farsighted decision-makers will in general be more inclined towards free trade than if they lack farsightedness. Further, we compared our results with those of the related literature in the simplest possible scenario, where only MTAs are permitted. Given that the differences in the simplest scenario are considerable, it should be expected that even greater differences will emerge in more complex scenarios, such as those involving situations where the decision-makers have access to all possible trade regimes.

In another contribution, Saggi and Yildiz (2010) show that when certain conditions are met, bilateral free trade is necessary for the achievement of global free trade, as a smaller country has a greater incentive to opt for global free trade under bilateral (with only FTAs available) rather than under multilateral agreements. Similarly, Lake (2017) demonstrates in a dynamic setup using the Markov Perfect Equilibrium solution concept that FTAs are "strong building blocs" (global free trade is only possible in the presence of FTAs) if the discount factor is sufficiently small. But when the discount factor exceeds a certain threshold, they become "strong stumbling blocs". In contrast to these results, when we consider a scenario in which only FTAs and MTAs are available, we find that FTAs never constitute 'building blocs'. Quite the contrary, in our paper, FTAs can only be viewed as 'stumbling blocs' in the setup similar to Saggi and Yildiz (2010) and Lake (2017) (see Figure 17 and compare scenarios noCUGFT and noPTAGFT). It is worth pointing out, however, that when both types of PTAs are included , the availability of FTAs may play an increasingly crucial strategic role (FTAs and CUs are strategic compliments) for CUs in maintaining the 'building bloc' nature of PTAs (see Figure 14 and Figure 17, region 6).

To conclude, we have demonstrated that complex strategic thinking coupled with farsightedness can provide more compelling explanations of which policy regimes are likely to be stable. As a result, we believe that the LCS solution concept is better suited to assess the stability of international trade policy regimes. Furthermore, we have highlighted the potential strategic complementarity between FTAs and CUs which, to the best of our knowledge, is novel to the literature.

5. Conclusion

Under the rules of the WTO (previously GATT), a group of countries can engage in both multilateral and different forms of preferential trade liberalization. The formation of global trade agreements is a complex game, and the rules of the game affect the exact nature of the outcome. Article I of the WTO aims at creating a multilateral free trading system, while Article XXIV allows countries to seemingly bypass the multilateral liberalization process. In this paper, our focus lies on the stability of trade policy arrangements under two different regulatory scenarios — with versus without PTAs. Based on the unlimited farsightedness of the participants in trade negotiations and considering an extensive set of trade agreements, we thus link our model more closely to reality.

Whether PTAs are 'building blocks' or 'stumbling blocks' on the path towards global free trade is not straightforward to answer. In the end, our results presented here provide a mixed answer, and we find that it depends on the size distribution of countries. Close to symmetry, global free trade is the unique stable trade constellation under both regulatory scenarios (PTA vs. no-PTA). By contrast, far away from symmetry, GFT might not be reached at all. In between, the effect of switching off Article XXIV depends on the exact form of asymmetry. In case two countries are relatively smaller than the third, prohibiting PTAs increases the area of stability of the GFT regimes. When two countries are relatively larger than the third, it reduces the area. Finally, when the world is far away from symmetry, abolishing the exception for PTAs might result in the worst possible state from the perspective of overall world welfare, the non-cooperative MFN regime. For such size distributions, PTAs act as a mechanism that helps to avoid the MFN regime.

Our research raises several interesting questions requiring further investigation. First, it will be interesting to study the robustness of our findings with respect to the underlying trade model. While the competing-exporters model has been the standard choice in the related literature, economists have also used both oligopoly and competing-importers models. The framework proposed here is general enough to accommodate different underlying trade models such as the ones just mentioned. Another potential area of inquiry might be an extension of the framework to allow a larger number of countries. In fact, in addition to bilateral negotiations, so-called plurilateral negotiations play an important role in the development of preferential trade liberalization. Recent examples include the Trans-Pacific Partnership (which has evolved into the CPTPP) and the Regional Comprehensive Economic Partnership (RCEP) championed by China. Including more than three countries in our model would allow us to investigate the strategic interactions among countries while taking such plurilateral negotiations into account. The incorporation of political economy considerations into the underlying trade model is another area of interest,[35] as it might allow us to understand the nature of tariff peaks occurring after PTAs come into effect. We believe that modifications and extensions of our framework (in line with these ideas) are directions worthy of further research.

As a final remark, it is perhaps essential, in the future, to lift the debate of 'building blocks' vs. 'stumbling blocks' to a level of detail that goes beyond such a binary view.

References

[1] General agreement on tariffs and trade, 1947. Source: `http://www.wto.org`.

[2] World trade report 2011, 2011. Source: `http://www.wto.org`.

[3] Philippe Aghion, Pol Antràs, and Elhanan Helpman. Negotiating free trade. *Journal of International Economics*, 73(1):1–30, 2007.

[4] Robert Aumann and Roger Myerson. Endogenous formation of links between players and coalitions: an application of the shapley value. *The Shapley Value*, pages 175–191, 1988.

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

[6] Kyle Bagwell and Robert W. Staiger. Regionalism and multilateral tariff cooperation. Working Paper 5921, National Bureau of Economic Research, February 1997.

[7] Kyle Bagwell and Robert W. Staiger. The design of trade agreements. *Handbook of Commercial Policy*, 1:435–529, 2016.

[8] Richard E. Baldwin and Thorvaldur Gylfason. A domino theory of regionalism. In *Expanding Membership of the European Union*, pages 25–53. Cambridge University Press, 1995.

[9] Douglas Bernheim, Bezalel Peleg, and Michael Whinston. Coalition-proof nash equilibria i. concepts. *Journal of Economic Theory*, 42(1):1–12, 1987.

[10] Jagdish Bhagwati. Regionalism and multilateralism: an overview. *New dimensions in regional integration*, 22:51, 1993.

[11] Michael Suk-Young Chwe. Farsighted coalitional stability. *Journal of Economic Theory*, 63(2):299 – 325, 1994.

[12] Giovanni Facchini, Peri Silva, and Gerald Willmann. The customs union issue: Why do we observe so few of them? *Journal of International Economics*, 90(1):136 – 147, 2013.

[13] Taiji Furusawa and Hideo Konishi. Free trade networks. *Journal of International Economics*, 72(2):310–335, 2007.

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

[15] Gene M Grossman. The purpose of trade agreements. *Handbook of Commercial Policy*, 1:379–434, 2016.

[16] John Harsanyi. An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Management Science*, 20(11):1472–1495, 1974.

[17] Jean-Jacques Herings, Ana Mauleon, and Vincent Vannetelbosch. Farsightedly stable networks. *Games and Economic Behavior*, 67(2):526–541, 2009.

[18] Jean-Jacques Herings, Ana Mauleon, and Vincent Vannetelbosch. Stability of networks under level-k farsightedness. Research Memorandum 030, Maastricht University, Graduate School of Business and Economics (GSBE), January 2014.

[35] See for example Facchini et al. (2013).

[19] Matthew Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

[20] Alexander Keck and Andreas Lendle. New evidence on preference utilization. WTO Staff Working Papers ERSD-2012-12, World Trade Organization (WTO), Economic Research and Statistics Division, September 2012.

[21] Hiau Looi Kee, Alessandro Nicita, and Marcelo Olarreaga. Estimating trade restrictiveness indices. *The Economic Journal*, 119(534):172–199, 2009.

[22] James Lake. Free trade agreements as dynamic farsighted networks. *Economic Inquiry*, 55(1):31–50, 2017.

[23] James Lake. Dynamic formation of preferential trade agreements: The role of flexibility. *Canadian Journal of Economics*, 52(1):132–177, 2019.

[24] Giovanni Maggi. International trade agreements. In *Handbook of International Economics*, volume 4, chapter 6, pages 317–390. 2014.

[25] Paul Missios, Kamal Saggi, and Halis Murat Yildiz. External trade diversion, exclusion incentives and the nature of preferential trade agreements. *Journal of International Economics*, 99:105–119, 2016.

[26] Debraj Ray and Rajiv Vohra. The farsighted stable set. *Econometrica*, 83(3):977–1011, 2015.

[27] Kamal Saggi, Alan Woodland, and Halis Murat Yildiz. On the relationship between preferential and multilateral trade liberalization: the case of customs unions. *American Economic Journal: Microeconomics*, 5(1):63–99, 2013.

[28] Kamal Saggi and Halis Murat Yildiz. Bilateralism, multilateralism, and the quest for global free trade. *Journal of International Economics*, 81(1):26–37, 2010.

[29] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[30] Jin Zhang, Zhiwei Cui, and Lei Zu. The evolution of free trade networks. *Journal of Economic Dynamics and Control*, 38:72–86, 2014.

[31] Jin Zhang, Licun Xue, and Lei Zu. Farsighted free trade networks. *International Journal of Game Theory*, 42(2):375–398, 2013.

Appendix A. Pseudocode

Note, that a couple of functions and variables are directly baked into the program without any further explanation in the pseudocode below - for example the matrix that determines the general network structure (for each player and all coalitions). The origin and characterization of these can be found in their respective parts in the main paper. The network structure A and the preference relations B both enter as a collection of $|X|\times|X|$-matrices, $\{A_S\}_{S\subseteq N,S\neq\emptyset}$ and $\{B_S\}_{S\subseteq N,S\neq\emptyset}$ resp., where $(A_S)_{i,j} = \mathbb{1}_{\{i\rightarrow_S j\}}(i,j)$ and $(B_S)_{i,j} = \mathbb{1}_{\{i\prec_S j\}}(i,j)$ for $(i,j)\in X\times X$.

Algorithm Largest Consistent Set

Input: Countries N, Outcomes X, Network Structure A, Preference Relations B
Output: Largest Consistent Set $\{Y\}$

procedure PARAMETERSPACELCS(N, X, A, B)
    $E = eMaxArea$ ▷ See Section 2.5
    $\alpha = \alpha MinValue(E)$ ▷ See Section 2.5
    for $e \in E$ **do**
        $Y = GeneralLCS(N, X, A, B)$
    return $\{Y\}$

function GENERALLCS(N, X, A, B)
    for $S \subseteq N$ **do**
        $C_S = \min\{A_S, B_S\}$
    $D^0 = \max_{S\subseteq N}\{C_S\}$ ▷ : Direct Dominance
    $n = 0$
    repeat
        $n = n + 1$
        for $S \subseteq N$ **do**
            $A^n_S = (\mathbb{1}_{\{(A_S\cdot D^{n-1})_{i,j}\neq 0\}}(i,j))_{(i,j)\in X\times X}$
            $D^n_S = \min\{A^n_S, B_S\}$
        $D^n = \max_{S\subseteq N}\{D^n_S\}$ ▷ : Indirect Dominance
    until $D^n = D^{n-1}$
    $D = \mathbb{1}_X + D^n$
    $Y^0 = (1)_{x\in X}$
    $m = 0$
    repeat
        $m = m + 1$
        for $x \in X$ **do**
            if $Y^{m-1}_x = 0$ **then**
                $Y^m_x = 0$
            else
                $y = \max\limits_{k\in X, S\subseteq N}\left\{(A_S)_{x,k}\left(1 - \max\limits_{z\in X}\left\{Y^{m-1}_z (D)_{k,z}\left(1-(B_S)_{x,z}\right)\right\}\right)\right\}$
                $Y^m_x = Y^{m-1}_x - y$
    until $Y^m = Y^{m-1}$
    $Y = Y^m$
    return Y

Appendix B. Model

B.1. **Individual Welfare.** The following table lists the individual welfare for each (representative) trade agreement, depending on endowments and tariffs, multiplied with the factor 18. Note that for MFN, CUGFT, FTAGFT, and MTAGFT the welfare W_i resembles W_j and W_k. In case of CU(i,j), FTA(i,j), and MTA(i,j) the welfare W_i is similar to W_j. For FTAHub(i) the welfare W_j resembles W_k.

Trade Agreement	Individual Welfare
MFN	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 - 8t_i^2 + t_j^2 + t_k^2 + 4e_i(9\alpha - e_j - e_k - t_j - t_k)$ $+2e_j(e_k + t_i + t_k) + 2e_k(t_i + t_j)$
CU(i,j)	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 - 11t_{ik}^2 + t_{jk}^2 + t_k^2 + 4e_i(9\alpha - e_j - e_k + t_{jk} - t_k)$ $+2e_j(e_k - 4t_{ik} + t_k) + 2e_k(5t_{ik} - t_{jk})$
↳ W_k	$2e_i^2 + 2e_j^2 - 10e_k^2 + 4t_{ik}^2 + 4t_{jk}^2 - 8t_k^2 + 2e_i(e_j - 2e_k + 2t_{jk} + t_k)$ $+2e_j(-2e_k + 2t_{ik} + t_k) + 4e_k(9\alpha - 2t_{ik} - 2t_{jk})$
CUGFT	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 + 4e_i(9\alpha - e_j - e_k) + 2e_j e_k$
FTA(i,j)	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 - 11t_{ik}^2 + t_{jk}^2 + t_k^2 + 4e_i(9\alpha - e_j - e_k + t_{jk} - t_k)$ $+2e_j(e_k - 4t_{ik} + t_k) + 2e_k(5t_{ik} - t_{jk})$
↳ W_k	$2e_i^2 + 2e_j^2 - 10e_k^2 + 4t_{ik}^2 + 4t_{jk}^2 - 8t_k 2 + 2e_i(e_j - 2e_k + 2t_{jk} + t_k)$ $+2e_j(-2e_k + 2t_{ik} + t_k) + 4e_k(9\alpha - 2t_{ik} - 2t_{jk})$
FTAHub(i)	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 + t_{jk}^2 + t_{kj}^2 + 4e_i(9\alpha - e_j - e_k + t_{jk} + t_{kj})$ $+2e_j(e_k - t_{kj}) - 2e_k t_{jk}$
↳ W_j	$2e_i^2 + 2e_j^2 - 10e_k^2 + 4t_{jk}^2 - 11t_{kj}^2 + 2e_i(e_j - 2e_k + 2t_{jk} - 4t_{kj})$ $+e_j(-4e_k + 10t_{kj}) + 4e_k(9\alpha - 2t_{jk})$
FTAGFT	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 + 4e_i(9\alpha - e_j - e_k) + 2e_j e_k$
MTA(i,j)	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 - 8t_i^2 + t_j^2 + t_k^2 + 4e_i(9\alpha - e_j - e_k - t_j - t_k)$ $+2e_j(e_k + t_i + t_k) + 2e_k(t_i + t_j)$
↳ W_k	$2e_i^2 + 2e_j^2 - 10e_k^2 + t_i^2 + t_j^2 - 8t_k^2 + 2e_i(e_j - 2e_k + t_j + t_k)$ $+2e_j(-2e_k + t_i + t_k) + 4e_k(9\alpha - t_i - t_j)$
MTAGFT	
↳ W_i	$-10e_i^2 + 2e_j^2 + 2e_k^2 + 4e_i(9\alpha - e_j - e_k) + 2e_j e_k$

Table 18. The individual welfare for each trade agreement depending on endowments and tariffs

B.2. **Tariffs.** The following describes the tariffs that the countries choose for each trade agreement. In addition to the specific restrictions mentioned in Section 2.1, all tariffs are bounded both from below and above by zero and the MFN-tariff respectively. As per WTO rule, the formation of any PTA does not allow additional tariffs towards others - which results in the upper bound of the MFN-tariff. Also, any form of subsidies is excluded here - which results in the lower bound of zero. Now, the following determines and describes the optimal tariffs for each scenario and the cases where capping occurs:

B.2.1. *MFN.* In this case, the optimal tariff of country i, given by $t_i^* = \frac{1}{8}(e_j + e_k)$, is always greater than zero as the endowments themselves are greater than zero. Additionally, t_i^* is going to play the role of the maximal tariff for country i for all the other agreements, then denoted t_i^{MFN}.

B.2.2. *CU.* Consider the scenario CU(i,j), then the optimal tariff of country i towards country k, given by $t_{ik}^* = \frac{1}{5}(2e_k - e_j)$, is always greater than zero but not always less than the MFN-tariff (and the one towards country j, t_{ij}^*, is always zero):

i) Lower Bound. By assumption on the endowments $e_k \geq \frac{3}{5}e_j$ and thus $e_k > \frac{1}{2}e_j$, which guarantees $t_{ik}^* > 0$.

ii) Upper Bound. By assumption on the endowments $e_k \leq \frac{5}{3}e_j$ however $t_{ik}^* \leq t_i^{MFN}$ requires $e_k \leq \frac{13}{11}e_j$, which leaves the interval $\frac{13}{11}e_j < e_k \leq \frac{5}{3}e_j$ to require capping. For this interval, the (maximal) MFN-tariff is optimal as the derivative of the joint welfare with respect to t_{ik} is always greater than zero on the interval $[0, t_i^{MFN}]$:

$$\frac{\partial(W_i + W_j)}{t_{ik}} = \frac{1}{9}\left(-10t_{ik} - 2e_j + 4e_k\right) \geq \frac{1}{36}\left(-13e_j + 11e_k\right) > 0$$

B.2.3. *FTA.* Consider the scenario FTA(i,j), then the optimal tariff of country i towards country k, given by $t_{ik}^* = \frac{1}{11}(5e_k - 4e_j)$, is neither always greater than zero nor always less than the MFN-tariff (but the one towards country j, t_{ij}^*, is zero):

i) Lower Bound. By assumption on the endowments $e_k \geq \frac{3}{5}e_j$ however $t_{ik}^* \geq 0$ requires $e_k \geq \frac{4}{5}e_j$, which leaves the interval $\frac{3}{5}e_j \leq e_k < \frac{4}{5}e_j$ to require capping. For this interval, the (minmal) zero-tariff is optimal as the derivative of the welfare with respect to t_{ik} is always lesser than zero on the interval $[0, t_i^{MFN}]$:

$$\frac{\partial W_i}{\partial t_{ik}} = \frac{1}{9}\left(-11t_{ik} - 4e_j + 5e_k\right) \leq \frac{1}{9}\left(5e_k - 4e_j\right) < 0$$

ii) Upper Bound. By assumption on the endowments $e_k \leq \frac{5}{3}e_j$ however $t_{ik}^* \leq t_i^{MFN}$ requires $e_k \leq \frac{43}{29}e_j$, which leaves the interval $\frac{43}{29}e_j < e_k \leq \frac{5}{3}e_j$ to require capping. For this interval, the (maximal) MFN-tariff is optimal as the derivative of the welfare with respect to t_{ik} is always greater than zero on the interval $[0, t_i^{MFN}]$:

$$\frac{\partial W_i}{\partial t_{ik}} = \frac{1}{9}\left(-11t_{ik} - 4e_j + 5e_k\right) \geq \frac{1}{72}\left(-43e_j + 29e_k\right) > 0$$

B.2.4. *MTA.* Consider the scenario MTA(i,j), then the optimal tariff of country i, given by $t_i^* = \frac{1}{7}(2e_k - e_j)$, is greater than zero and less or equal to the MFN-tariff as per assumption on the endowments $\frac{3}{5}e_j \leq e_k \leq \frac{5}{3}e_j$.

B.2.5. *Notes.* The analysis considered country i and an agreement with country j, but it naturally extends to all other combinations. Also, the perspective of the third country needs no further analysis as it always chooses the MFN-tariff. Furthermore, the case of FTAHub(i) is simply a combination of FTA(i,j) and FTA(i,k). Finally, the three variants of GFT require no additional analysis as every country always chooses the zero-tariff.

B.3. **Overall Welfare.** The following table lists the overall welfare for each (representative) trade agreement, depending purely on endowments, computed modulo $2\alpha\left(\sum_{n\in N} e_n\right)$, which is the common term associated with the factor α. Also, the notation l_c and l^c is used to indicate that country l is capped in terms of tariffs from below or above respectively.

Trade Agreement	Overall Welfare
MFN	
↳ no cap	$\frac{11}{32}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
CU(i,j)	
↳ no cap	$\frac{1}{1600}(-563e_i^2 - 550e_ie_j - 448e_ie_k - 563e_j^2 - 448e_je_k - 704e_k^2)$
↳ i^c	$\frac{1}{1600}(-563e_i^2 - 550e_ie_j - 448e_ie_k - 550e_j^2 - 550e_je_k - 627e_k^2)$
↳ i^c, j^c	$\frac{11}{32}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
CUGFT	
↳ no cap	$\frac{1}{3}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
FTA(i,j)	
↳ no cap	$\frac{1}{7744}(-2963e_i^2 - 2662e_ie_j - 1728e_ie_k - 2963e_j^2 - 1728e_je_k - 3648e_k^2)$
↳ i_c	$\frac{1}{23232}(-8889e_i^2 - 7986e_ie_j - 5184e_ie_k - 7865e_j^2 - 7744e_je_k - 9344e_k^2)$
↳ i_c, i_c	$\frac{1}{192}(-65e_i^2 - 66e_ie_j - 64e_ie_k - 65e_j^2 - 64e_je_k - 64e_k^2)$
↳ i^c	$\frac{1}{7744}(-2963e_i^2 - 2662e_ie_j - 1728e_ie_k - 2662e_j^2 - 2662e_je_k - 3155e_k^2)$
↳ i^c, j^c	$\frac{11}{32}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
FTAHub(i)	
↳ no cap	$\frac{1}{363}(-153e_i^2 - 81e_ie_j - 81e_ie_k - 146e_j^2 - 121e_je_k - 146e_k^2)$
↳ j_c	$\frac{1}{363}(-137e_i^2 - 81e_ie_j - 121e_ie_k - 146e_j^2 - 121e_je_k - 121e_k^2)$
↳ j_c, k_c	$\frac{1}{3}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
↳ j^c	$\frac{1}{23232}(-8889e_i^2 - 5184e_ie_j - 7986e_ie_k - 9344e_j^2 - 7744e_je_k - 7865e_k^2)$
↳ j^c, k^c	$\frac{1}{192}(-66e_i^2 - 66e_ie_j - 66e_ie_k - 65e_j^2 - 64e_je_k - 65e_k^2)$
FTAGFT	
↳ no cap	$\frac{1}{3}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$
MTA(i,j)	
↳ no cap	$\frac{1}{3136}(-1083e_i^2 - 1078e_ie_j - 960e_ie_k - 1083e_j^2 - 960e_je_k - 1216e_k^2)$
MTAGFT	
↳ no cap	$\frac{1}{3}(-e_i^2 - e_ie_j - e_ie_k - e_j^2 - e_je_k - e_k^2)$

TABLE 19. The overall welfare for each trade agreement depending on endowments

C.1. **Proofs of propositions.**

C.1.1. *Proof of the proposition 5.*

Proof. Let us start by giving the indirect dominance matrix:

$$
\begin{array}{l}
\\
1\,MFN \\
2\,CU(a,b) \\
3\,CU(b,c) \\
4\,CU(c,a) \\
5\,CUGFT \\
6\,FTA(a,b) \\
7\,FTA(b,c) \\
8\,FTA(c,a) \\
9\,FTAHub(a) \\
10\,FTAHub(b) \\
11\,FTAHub(c) \\
12\,FTAGFT \\
13\,MTA(a,b) \\
14\,MTA(b,c) \\
15\,MTA(c,a) \\
16\,MTAGFT
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16
\end{array} \\
\begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0
\end{pmatrix}
\end{array}
$$

Recall that X denotes the full set and let $Y = \{CU(b,c), FTA(b,c)\}$ be the candidate for the LCS. Take any element x from the set $X \setminus Y$ and consider the deviation $x \rightarrow_{\{b,c\}} CU(b,c)$. Note that $CU(b,c)$ is not indirectly dominated by any other element from X and furthermore $x \prec_{\{b,c\}} CU(b,c)$ for all $x \in X \setminus Y$. Thus, the deviation $x \rightarrow_{\{b,c\}} CU(b,c)$ can not be deterred for all $x \in X \setminus Y$. Therefore, no such x can be part of the stable set.[36]

As each outcome in $X \setminus Y$ is indirectly dominated by $y \in Y$ (see the matrix), for any coalition and any deviation away from $y \in Y$ there always exists a path of indirect dominance back to Y. Moreover, no coalition is actually better off when coming back to Y, as $x \not\prec_S y$ for all $x, y \in Y$, $x \neq y$, and $S \subseteq N$, $S \neq \emptyset$. Therefore, the set Y satisfies the (internal) stability condition while being maximal, i.e. $Y = LCS$. □

C.1.2. *Proof of the proposition 6.*

Proof. The indirect dominance matrix is given as follows:

[36] It might appear that this proof deviates from the general approach of eliminating element by element from the full set until the remainder forms the stable set. However, in this proof it is purely a coincidence that in one step all elements but the stable ones can be eliminated with one argument (or rather deviation).

$$
\begin{array}{l} \\ 1\,MFN \\ 2\,MTA(a,b) \\ 3\,MTA(b,c) \\ 4\,MTA(c,a) \\ 5\,MTAGFT \end{array}
\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0
\end{pmatrix} \end{array}
$$

Start with the full set again. If we consider the deviations $MTA(c,a) \to_c MFN$ and $MTA(a,b) \to_b MFN$, then no further deviations are expected as MFN is not indirectly dominated by any other outcome. In addition, $MTA(c,a) \prec_c MFN$ and $MTA(a,b) \prec_b MFN$, so $MTA(c,a)$ and $MTA(a,b)$ cannot be part of the stable set. The same argument works in the case of $MTAGFT$ and the deviation $MTAGFT \to_{b,c} MFN$, as $MTAGFT \prec_{b,c} MFN$. So, the global free trade regime cannot be stable as well.

Let $Y = \{MFN, MTA(b,c)\}$. Following any deviation from the elements in Y, there is always an indirect dominance path coming back to Y (MFN in this case). In addition, for any $x, y \in Y$ with $x \neq y$ there does not exist coalition S for which $x \prec_S y$. Thus, the set Y is consistent and the largest one as well. □

C.1.3. *Proof of the proposition 7.*

Proof. The indirect dominance matrix is given as follows:

$$
\begin{array}{l} \\ 1\,MFN \\ 2\,CU(a,b) \\ 3\,CU(b,c) \\ 4\,CU(c,a) \\ 5\,CUGFT \\ 6\,FTA(a,b) \\ 7\,FTA(b,c) \\ 8\,FTA(c,a) \\ 9\,FTAHub(a) \\ 10\,FTAHub(b) \\ 11\,FTAHub(c) \\ 12\,FTAGFT \\ 13\,MTA(a,b) \\ 14\,MTA(b,c) \\ 15\,MTA(c,a) \\ 16\,MTAGFT \end{array}
\begin{array}{c} \begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix} \end{array}
$$

First, take $x \in \{CU(i,b), FTA(i,b), MTA(i,b), FTAHub(i), FTAHub(b)\}$, with $i \in \{a,c\}$. Country b can destroy such trade agreements and, depending on the initial constellation, either $FTA(c,a)$ or the MFN regime remains. Then, further deviations are possible, namely $MTA(c,a)$ and $CU(c,a)$. However, each of the aforementioned trade agreements is indirectly dominated by $CU(c,a)$ and simultaneously country b is better off compared to the initial situation. Consequently, such deviations can not be avoided and no such x can be part of the stable set.

Now, consider $x \in \{MFN, FTA(c,a), MTA(c,a)\}$ for which $x \to_{\{a,c\}} CU(a,c)$ presents a deviation that can not be deterred. As in the previous paragraph, $CU(c,a)$ is not indirectly dominated any element and also $x \prec_{\{a,c\}} CU(a,c)$. Thus, no such x can be the part of the stable set as well.

At last, let $x \in \{CUGFT, FTAGFT, MTAGFT\}$ and consider the deviations where country b leaves the agreements. $CU(c,a)$, $FTA(c,a)$, or $MTA(c,a)$ can be the result. We have shown that the last two outcomes can not be stable. As for $CU(a,c)$, we have that for all x considered $x \prec_{\{b\}} CU(a,c)$. As a result, we conclude that no such x can be in the consistent set.

$CU(a,c)$ indirectly dominates each outcome, all deviations from it are deterred. So, the set containing $CU(a,c)$ is consistent and the largest one as well. $\square$

C.1.4. *Proof of the proposition 8.*

Proof. In this case, the indirect dominance matrix has the following form:

$$
\begin{array}{l}
 \\
1\,MFN \\
2\,MTA(a,b) \\
3\,MTA(b,c) \\
4\,MTA(c,a) \\
5\,MTAGFT
\end{array}
\begin{array}{c}
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

Assume, $x \in \{MTA(a,b), MTA(b,c), MTAGFT\}$ and consider the deviations, where country b dismantles any above mentioned constellation. Two possibilities: Either MFN or $MTA(c,a)$ remain. From MFN either no coalition moves away or, as it is indirectly dominated by $MTA(c,a)$ (see the indirect dominance matrix), the latter might be approached. In either case, b is better off. Thus, no such x can be part of the stable set.

Now, analyse the case of the MFN regime. Take the following deviation: $MFN \to_{\{a,c\}} MTA(c,a)$. As $MTA(c,a)$ is not indirectly dominated by any other trade agreement and $MFN \prec_{\{a,c\}} MTA(c,a)$, the MFN regime can not be stable as well.

As $MTA(c,a)$ indirectly dominates each trade agreement, all deviations from it are deterred. So, the set consisting of $MTA(c,a)$ is consistent and the largest one as well. $\square$

C.2. **Exact Intervals.** The tables here list the exact intervals within which each specific trade agreement is stable (for the boundaries of the parameter space).:

Trade Agreement	Exact Interval(s)
$e_b = e_{\min} \leq e_c \leq e_{\max} = e_a$	
CU(b,c)	[1.00000000000000,1.33800935203741]
CU(c,a)	[1.32598530394122,1.66666666666667]
FTA(b,c)	[1.00000000000000,1.38076152304609]
FTA(c,a)	[1.33533734134937,1.63059452237809]
FTAHub(c)	[1.23647294589178,1.26987307949232]
	[1.33533734134937,1.36472945891784]
MTA(c,a)	[1.3794255177021,1.6359385437542]
$e_b = e_c = e_{\min} \leq e_a \leq e_{\max}$	
CU(a,b)	[1.02404809619238,1.11088844355377]
CU(b,c)	[1.02404809619238,1.66666666666667]
CU(c,a)	[1.02404809619238,1.11088844355377]
CUGFT	[1.00000000000000,1.18036072144289]
FTA(a,b)	[1.02404809619238,1.24048096192385]
FTA(b,c)	[1.02404809619238,1.291249164996]
	[1.4836339345357,1.66666666666667]
FTA(c,a)	[1.02404809619238,1.24048096192385]
FTAHub(b)	[1.00133600534402,1.291249164996]
FTAHub(c)	[1.00133600534402,1.291249164996]
FTAGFT	[1.00000000000000,1.18036072144289]
MTA(a,b)	[1.02404809619239,1.14696058784235]
MTA(c,a)	[1.02404809619239,1.14696058784235]
MTAGFT	[1.00000000000000,1.18036072144289]
$e_b = e_{\min} \leq e_a = e_c \leq e_{\max}$	
MFN	[1.20440881763527,1.22979291917168]
	[1.27521710086840,1.2765531062124]
CU(a,b)	[1.04542418169673,1.24983299933199]
CU(b,c)	[1.04542418169673,1.24983299933199]
CU(c,a)	[1.04943219772879,1.66666666666667]
CUGFT	[1.00000000000000,1.22979291917168]
FTA(a,b)	[1.04943219772879,1.29258517034068]
FTA(b,c)	[1.04943219772879,1.29258517034068]
FTA(c,a)	[1.04542418169673,1.29258517034068]
FTAHub(a)	[1.04542418169673,1.2765531062124]
FTAHub(b)	[1.04542418169673,1.29258517034068]
FTAHub(c)	[1.04542418169673,1.2765531062124]
FTAGFT	[1.00000000000000,1.22979291917168]
MTA(a,b)	[1.04943219772879,1.22444889779559]
MTA(b,c)	[1.04943219772879,1.22444889779559]
MTA(c,a)	[1.04542418169673,1.29258517034068]
MTAGFT	[1.00000000000000,1.22979291917167]

TABLE 20. The exact intervals of stability with PTAs

Trade Agreement	Exact Interval(s)
	$e_b = e_{\min} \le e_c \le e_{\max} = e_a$
MFN	[1.0000000000000000,1.3780895123580494]
MTA(b,c)	[1.0000000000000000,1.0000000000000000]
MTA(c,a)	[1.379425517702071,1.6666666666666667]
	$e_b = e_c = e_{\min} \le e_a \le e_{\max}$
MFN	[1.6666666666666667,1.6666666666666667]
MTA(a,b)	[1.0307281229124916,1.1469605878423514]
MTA(b,c)	[1.0307281229124916,1.1469605878423514] [1.3754175016700068,1.6666666666666667]
MTA(c,a)	[1.0307281229124916,1.1469605878423514]
MTAGFT	[1.0000000000000000,1.3740814963259853]
	$e_b = e_{\min} \le e_a = e_c \le e_{\max}$
MTA(a,b)	[1.0160320641282565,1.1202404809619237]
MTA(b,c)	[1.0160320641282565,1.1202404809619237]
MTA(c,a)	[1.0160320641282565,1.6666666666666667]
MTAGFT	[1.0000000000000000,1.203072812291249]

TABLE 21. The exact intervals of stability without PTAs

Trade Agreement	Exact Interval(s)
	$e_b = e_{\min} \le e_c \le e_{\max} = e_a$
CU(b,c)	[1.000000000000,1.338009352037]
CU(c,a)	[1.325985303941,1.666666666667]
MTA(c,a)	[1.37942551770,1.63593854375]
	$e_b = e_c = e_{\min} \le e_a \le e_{\max}$
CU(b,c)	[1.024048096192,1.666666666667]
CUGFT	[1.000000000000,1.02404809619238]
MTAGFT	[1.000000000000,1.02404809619238]
	$e_b = e_{\min} \le e_a = e_c \le e_{\max}$
CU(a,b)	[1.045424181697,1.26052104208417]
CU(b,c)	[1.045424181697,1.26052104208417]
CU(c,a)	[1.049432197729,1.666666666667]
CUGFT	[1.000000000000,1.229792919172]
MTA(a,b)	[1.049432197729,1.224448897796]
MTA(b,c)	[1.049432197729,1.224448897796]
MTA(c,a)	[1.045424181697,1.26052104208417]
MTAGFT	[1.000000000000,1.229792919172]

TABLE 22. The exact intervals of stability without FTAs

Trade Agreement	Exact Interval(s)
$e_b = e_{\min} \leq e_c \leq e_{\max} = e_a$	
FTA(b,c)	[1.00000000000000,1.38076152304609]
FTA(c,a)	[1.33533734134937,1.63059452237809]
FTAHub(c)	[1.23647294589178,1.26987307949232]
	[1.33533734134937,1.36472945891784]
MTA(c,a)	[1.37942551770207,1.66666666666667]
$e_b = e_c = e_{\min} \leq e_a \leq e_{\max}$	
FTA(a,b)	[1.00267201068804,1.24048096192385]
FTA(b,c)	[1.00267201068804,1.66666666666667]
FTA(c,a)	[1.00267201068804,1.24048096192385]
FTAHub(a)	[1.062792251169,1.13627254509018]
FTAHub(b)	[1.00267201068804,1.32464929859719]
FTAHub(c)	[1.00267201068804,1.32464929859719]
FTAGFT	[1.00000000000000,1.18036072144289]
MTA(a,b)	[1.00267201068804,1.14696058784235]
MTA(b,c)	[1.00267201068804,1.13627254509018]
MTA(c,a)	[1.00267201068804,1.14696058784235]
MTAGFT	[1.00000000000000,1.18036072144289]
$e_b = e_{\min} \leq e_a = e_c \leq e_{\max}$	
FTA(a,b)	[1.00133600534402,1.27388109552438]
FTA(b,c)	[1.00133600534402,1.27388109552438]
FTA(c,a)	[1.00133600534402,1.27388109552438]
FTAHub(a)	[1.05076820307281,1.27388109552438]
FTAHub(b)	[1.00133600534402,1.04943219772879]
	[1.12024048096192,1.27388109552438]
FTAHub(c)	[1.05076820307281,1.27388109552438]
FTAGFT	[1.00000000000000,1.20307281229125]
MTA(a,b)	[1.00133600534402,1.08550434201737]
	[1.09218436873748,1.12024048096192]
MTA(b,c)	[1.00133600534402,1.08550434201737]
	[1.09218436873748,1.12024048096192]
MTA(c,a)	[1.09218436873748,1.66666666666667]
MTAGFT	[1.00000000000000,1.20307281229125]

TABLE 23. The exact intervals of stability without CUs