Bilateralism versus Multilateralism: Role of Economies of Scale for the Prospect of Global Free Trade

Muhammad Azeem * Halis Murat Yildiz†

November 15, 2025

Abstract

In this paper, we examine how the freedom to pursue free trade agreements (FTAs) affects the external tariffs, welfare of countries and the prospect of global free trade relative to MFN-based trade liberalization when industry exhibits flexible degrees of economies of scale in an imperfectly competitive environment. We show that both the degree of asymmetry in market sizes across countries and the degree of economies of scale are pivotal in determining: (i) the external tariffs of both member and non-member countries; (ii) the strength of free riding incentive of the non-member country and exclusion incentives of member countries, and (iii) whether the ability to form FTAs acts as building or stumbling blocks towards global free trade. We first show that, under both bilateralism and multilateralism games, global Free Trade is the unique CPNE when countries are relatively symmetric in market size and such an outcome is more likely as the degree of economies of scale rises. We then show, when countries are sufficiently asymmetric in market size, that the ability to form bilateral FTAs can be necessary to reach global free trade: global free trade can fail to arise as CPNE under MFN based game while it is the unique CPNE under the FTA game. Over this range, we argue that the ability to form bilateral FTAs acts a strong building block for the prospect of multilateral free trade and this outcome becomes more likely as the degree of economies of scale rises.

Keywords: Economies of Scale, Tariff Complementarity, Free Trade Agreement, Preferential Trade Agreement, Coalition Proof Nash Equilibrium

JEL Classifications: C72, F12, F13.

*Department of Economics, Toronto Metropolitan University. Email: ma122@torontomu.ca
†Department of Economics, Toronto Metropolitan University. Email: hyildiz@torontomu.ca

1 Introduction:

Since the establishment of the General Agreements on Tariffs and Trade (GATT), now subsumed by the World Trade Organization (WTO), trade liberalization proceeded primarily along two fronts: periodic rounds of multilateral negotiations which yield Most Favored Nation (MFN) tariff concessions in a non-discriminatory fashion, and the formation of Preferential Trade Agreements (PTAs) that facilitate discriminatory trade liberalization amongst a subset of WTO members, as permitted under GATT Article XXIV. Historically, multilateral trade liberalization has led to substantial MFN tariff concessions, with countries undertaking binding commitments to limit their maximum applied MFN tariffs. While being effective historically, multilateral negotiations have stalled recently. Doha round of negotiations failed to produce an acceptable agreement for all countries despite extensive negotiations. In contrast, PTAs has surged in popularity. Following the completion of the Uruguay Round in 1994, the number of PTAs has grown exponentially, with over 500 currently in force. Roughly 90 percent of these agreements are Free Trade Agreements (FTAs), with the rest predominantly consisting of Customs Unions (CUs).

The rapid proliferation of PTAs has raised serious concerns among economists and policy-makers about its potential to undermine the ultimate objective of attaining global free trade. The discussion frequently focuses on how the PTA formation affects the external tariffs on the non-members and whether the formation of PTAs act as building or stumbling blocs towards global free trade, compared to the results that would arise with pure multilateralism. [1]

In general, the optimal external tariffs of an FTA member country are largely determined by how the formation of an FTA affects trade flows. Changes in trade flows, in turn, are highly sensitive to the returns to scale in the production technology. Depending on whether production technology exhibits economies of scale, diseconomies of scale or constant returns to scale, the trade flows and, thus, the external tariffs of member countries are affected. In most of the literature, it is customary to assume the role of different degrees of economies of scale away and this imposes a serious limitation on the analysis and may explain the ambiguity in the existing

[1] Notable contributions to this body of literature, exploring the implications of PTAs on optimal tariffs, reduction of multilateral tariffs, and their impact on achieving global free trade, include studies by Krugman (1991), Bhagwati (1991), Yi (1996), Bagwell and Staiger (1997a, b, 1998, 2005a, b), Krishna (1998), Riezman (1999), Bond et al. (2004), Bond (2009), Goyal and Joshi (2006), Furusawa and Konishi (2007), Aghion et al. (2007), Missios et al. (2016), Ornelas (2005, 2007), Saggi (2006), Saggi and Yildiz (2010, 2011), Saggi et al. (2013, 2018, 2019), Stoyanov and Yildiz (2015), and Lake (2017), Lake et al. (2020), Chi and Yildiz (2024a, 2024b).

empirical studies on the effect of PTAs on trade, tariffs and welfare.

The empirical evidence on average industry returns to scale is mixed with a consensus that the estimates are heterogeneous across industries and highly sensitive to the level of aggregation. Assessing returns to scale has been a long and contentious debate surrounding the proper identification of production functions. In earlier works, Hall (1988, 1990) finds large and significant deviations from constant returns to scale using instrumental variables estimation with value-added data, pointing to quantitatively significant increasing returns to scale in two digit US manufacturing industry. [2] Using a general equilibrium model of international trade and covering 27 manufacturing and 7 non-manufacturing industries, Antweiler and Trefler (2002) obtained a significant deviation from CRS, especially for a third of industries (including pharmaceuticals and machinery). Recently, researchers in industrial organization, such as Levinsohn and Petrin (2003), Ackerberg et al. (2015), Wooldridge (2009), Gandhi et al. (2013), and Grieco et al. (2016) have developed structural estimation techniques that overcome the endogeneity and selection problems in the estimation of production functions. Gao and Kehrig (2017) and Bartelme et al. (2024) estimate the economies of scale using these various modern structural estimation techniques and found large differences in returns to scale across industries. The differences in returns to scale are remarkable and shape the cost structure, production, trade flows and trade patterns. Therefore, they have significant implications on trade policy and welfare implications of FTAs. Motivated by these facts, we allow for production technologies to exhibit different degrees of economies of scale across sectors in an imperfectly competitive environment. To the best of our knowledge, this paper is the first paper to incorporate flexible degree of economies of scale in examining the external tariff setting of FTA member countries and the equilibrium FTA formation.

We first show that the existence of economies of scale or diseconomies of scale creates a link between outputs of a firm in the domestic and export markets, and thus, unlike the constant returns to scale, markets are not strategically separated. Our first finding suggests that the tariff complementarity effect always exists regardless of the degree of economies of scale. The intuition behind this results is that, due to the preferential access of members in each others' markets and the discrimination non-member firms face in the members' markets, the non-members become less important rent extraction source for members, triggering the incentives of

[2] Basu and Fernald (1997) and Burnside (1996) note that estimates of returns to scale are very sensitive to the levels of aggregation, often rising at higher levels of aggregation and vary widely across relatively disaggregated industries.

member countries to impose lower external tariffs. Then we go one step further and argue that the degree of tariff complementarity rises as the degree of economies scale get stronger. This is primarily due to the fact that, as the degree of economies of scale gets stronger, PTA formation leads to higher efficiency gains for the member countries' firms, and thus the market share of member countries's firms rises even more, making the non-member country even less important source of rent extraction. The existence of tariff complementarity has been supported bt the empirical studies by Estevadeordal et al. (2008) Calvo-Pardo et al. (2009) and Bohara et al. (2004). On the other hand, Limão (2006) argues the opposite in his empirical study, by looking at US tariffs. Our results provide additional insights on these mixed empirical findings, arguing that the degree of tariff complementarity effect critically depends upon the degree of economies of scale. This is a critical observation and provides us a testable hypothesis that is carried out in our companion paper Azeem et al. (2024a).

Having established the fact that the external tariffs of member countries are largely affected by different degrees of economies of scale, we next move to the external tariff setting in the non-member country. We know from the competing importers literature (see Missios et al. 2016; Saggi et al. 2018 and Chi and Yildiz (2022)) that FTA formation leads to external trade diversion: export supply elasticity of member countries rise due to PTA formation, lowering the external tariffs of the non-member countries. Note that this literature focuses on decreasing returns to scale in a perfectly competitive environment. In the present paper, we take this discussion further to a setting where we have flexible degrees of economies of scale. Our findings support the above mentioned existing literature on the external trade diversion only when the firms operate with diseconomies of scale. However, when there are economies of scale, we find the opposite result. As discussed above, under the economies of scale, PTA formation leads to higher mark-up for the firms in member countries, raising the incentive of the non-member country to extract rent via higher external tariffs. In Azeem et al. (2024b), empirical testing provides a strong support for this result.

Next, we move to the endogenous formation of FTAs. The debate on the implications of the formation of PTAs on the prospect of multilateral free trade centers around the terms building and stumbling blocks, introduced by Bhagwati (1991). The standard view in the existing literature is that PTAs are building (stumbling) blocks to global free trade if the process of preferential trade liberalization eventually converges (fails to converge) to global free trade.[3]

[3]Grossman and Helpman ()1995) Krishna (1998) and Ornelas (2005a,b) treat trade agreements exogenously

However, as highlighted by Saggi and Yildiz (2010), Saggi et al. (2013) and Stoyanov and Yildiz (2015), this view can be misleading if global free trade obtains regardless of whether the formation of PTAs is allowed or not. In this paper, we also take this alternative view. There is a large extant literature on the PTA formation and its implications on the prospects of multilateral free trade in a variety of setting (see Maggi 2014 for an excellent survey). The theoretical and empirical results are mixed depending on the degree of asymmetry with respect to market size, technology, endowments and political economy factors (see Aghion et al., 2007; Konishi and Furusawa (2007; Saggi and Yildiz, 2010 and Stoyanov and Yildiz, 2015). Specifically, focusing on the FTA formation incentives, Saggi (2006), Saggi and Yildiz (2011), Ornelas (2005), and Stoyanov and Yildiz (2015) employed a similar oligopoly model of trade as in the present paper with a CRS assumption and thus ignore the role different degrees of economies of scale play in the PTA formation. Since the welfare properties of FTAs and thus the incentives of countries to form FTAs heavily rely on the scale economies of production technology, it is crucial to understand the equilibrium path of PTAs and whether FTA formation undermines multilateral trade liberalization in a general setting.

We consider a world in which all countries are the WTO members and apply the MFN tariff rate towards each other unless they form a preferential trade agreement. A group of countries can choose to lower tariffs towards each other as long at these preferences are extended to other countries, [4] and we call such an arrangement an MFN based agreement. We next consider Article XXIV world where countries have the option to liberalize trade on preferential basis and form Free Trade Agreements (FTAs). Our approach to trade agreement formation follows that of Saggi and Yildiz (2010) under which each country is free to form bilateral FTAs, MFN based agreement or pursue no trade liberalization. Incorporating flexible degrees of economies of scale with asymmetric market size, we take the building vs. stumbling blocks debate one step further. To this end, we employ a three-stage game under both FTA and MFN based trade liberalization games. In the first stage of each game, equilibrium trade agreements are formed between countries. In the second stage, countries choose import tariffs under a given trade policy regime, and in the third stage firms compete in a Cournot fashion. Using the concept of coalition-proof Nash equilibrium (CPNE), we identify the consequences of the exception to

while Goyal and Joshi (2006), Aghion et al. (2007), Furusawa and Konishi (2007), and Seidman (2009) use endogenous formation of PTAs but ignore the possibility of trade liberalization on an MFN basis, and view PTAs as building blocs so long as their pursuit eventually leads to global free trade.

[4] We do not consider unilateral trade liberalization since in our model all countries have unilateral incentives to extract tariff rent and such liberalization is not in any country's interest.

non-discrimination principle allowed by the GATT Article XXIV via comparing the CPNE agreements under the FTA and MFN based games. Saggi and Yildiz (2010) show that the option to form FTAs acts as building blocs only there exist circumstances where global free trade is an equilibrium only if countries are free to form bilateral FTAs. When such is the case, we say that FTAs act as strong building blocks. Our findings in this paper support these results even with flexible degrees of economies of scale. Note that, while Saggi and Yildiz (2010) uses an endowment model with CRS production technology, we allow for a flexible degree of economies of scale in an imperfectly competitive environment where the market size across countries are potentially asymmetric.

We first derive the Nash equilibria for both the FTA and MFN-based games and show that multiplicity problem arises. To address this issue, as in Saggi and Yildiz (2010), Saggi et al. (2013), Stoyanov and Yildiz (2015), and Chi and Yildiz (2022), we refine the Nash equilibria using the concept of Coalition-Proof Nash Equilibrium (CPNE) that is immune to self-enforcing coalitional deviations. We then compare the CPNE of the FTA game with those under the MFN-based game. This comparison provides an analysis of whether the ability to form FTAs can serve as stumbling blocks or building blocks, depending on the relative strength of economies of scale and market size asymmetry.

Before proceeding with the equilibrium analysis in the FTA game, we should note that global free trade may fail to be a CPNE due to two reasons: (i) FTA member countries' may exclude the third country from the free trade network, called as the 'exclusion incentive' or (ii) a non-member country's may remain as an outsider under an FTA and free-ride on tariff reductions by the FTA member countries, called as the 'free-riding incentive'.[5] We first show that, while there exist exclusion incentives for countries with medium and large market sizes, these incentives go unexercised due to the flexibility of FTA formation since one of the inititially excluding countries has an incentive to deviate further to become the hub country under a hub and spoke regime. Therefore, our focus will be on the free riding incentives. On one hand, the samller country always benefit from an FTA with the larger country unless it becomes a spoke country. Therefore, the smaller country has no incentive to unilaterally deviate from free trade. On the other hand, medium and large countries have incentives to break up an agreement if and only if they are sufficiently large relative to their smaller partners. These findings together

[5]Saggi and Yildiz (2010), Saggi et al. (2013), Lake and Krishna (2019), and Chi and Yildiz (2022) discuss the role of exclusion incentives and free-riding incentives in attaining global free trade.

imply that the unilateral incentive of a large country to unilaterally deviate from free trade (and thus free ride) determines the CPNE condition of free trade. Given the above discussion, we first show that global free trade is the unique CPNE regardless of the degrees of economies of scale when countries are sufficiently symmetric in market size. Under such a case, there exists no self-enforcing deviation from free trade. This is a confirmation that relatively symmetric countries always have incentives to exchange preferential market access reciprocally and free trade obtains. However, as the degree of market size asymmetry rises between countries, the large country has less incentive to engage in free trade negotiation since it becomes costly to give up its tariff ability in a larger market in exchange for preferential access to a smaller market. However, note also that as the degree of economies of scale rises, the objection of large country to FTA formation gets weaker. Under such a case, our first main result emerges: when countries are moderately asymmetric, global free trade only arises as a CPNE when the degree of economies of scale is sufficiently high. Note that the higher the degree of economies of scale, the larger the average cost reduction following a formation of an FTA via higher volume of trade and this efficiency gain convinces the large country no deviate from global free trade. When the degree of market size asymmetry exceeds a certain treshold, global free trade fails to be an equilibrium regardless of the degree of economies of scale. Under such a case, large country has no interest in engaging to free trade negotiations and the FTA between small and medium countries obtains as the CPNE. When the market size asymmetry is extremely large, then no agreement obtains as the unique CPNE.

Next, we move to the MFN-based game where countries can liberalize only on MFN basis. Under a two country MFN-based agreement, two insiders jointly maximize their payoffs with the constraint that they must extend the same tariff concession to the third country. Our findings indicate that countries with small and medium market sizes have no incentives to unilaterally or jointly deviate from free trade. As in the FTA game, the CPNE condition for global free trade depends solely on the free riding incentive of the large country that falls with the degree of economies of scale. This implies that global free trade is more likely to be a CPNE when industry exhibits stronger economies of scale. In order to highlight whether the ability to form an FTA affects the prospect of global free trade, we compare the free-riding incentives in the FTA and MFN-based games. Note that while the large country faces lower tariffs in the FTA between small and medium countries relative to two country MFN-based agreement, it comes in a discriminatory fashion and the discrimination effect dominates the tariff level effect, making

the free riding incentives under the MFN-based agreement stronger. Therefore, we argue that when the market size asymmetry is in the medium range, there is a case where Article XXIV is necessary for achieving global free trade that would not have been obtained if countries are forced to liberalize trade only on MFN basis. Under such a case, we argue that FTAs act as a strong building block. Note that this result is more likely when the degree of economies of scale is higher.

The paper proceeds as follows. The next section presents the underlying trade model and the optimal tariff-setting. Section 3 outlines the strategies and equilibrium concepts along with their stability (CPNE), with a particular focus on the welfare effects of FTAs. Section 4 presents the Endogenous MFN-based Agreement Formation game. Finally, Section 5 provides a comparison between both games, and Section 6 concludes the discussion.

2 Model:

We utilize an adapted version of Brander and Krugman 1983 oligopoly trade model. There are three symmetric countries: i, j, k and two goods: a non-numeriare good x and a numeriare good y. Hereafter, due to symmetry, we use country i as a generic notation for countries.

On the demand side, the representative consumer's utility function is quasi-linear

$$U_i\left(x_i, y_i\right) = u_i\left(x_i\right) + y_i$$

where the subutility $u_i(x_i)$ is quadratic: $u_i(x_i) = A_i x_i - \frac{x_i^2}{2}$ where x_i is the total output sold in country i: $x_i = \sum_{z=i,j,k} x_{zi}$. The demand function for good x is linear in each country: $p_i = A_i - x_i$, where p_i denotes the price of good x_i in country i. Note that x_{zi} denotes the sale of country z's firm in country i. In order to highlight the degree of asymmetry in market size, we let $A_s = 1 - \alpha$, $A_m = 1$, and $A_\ell = 1 + \alpha$. Here, note that α represents the degree of market size asymmetry both between countries s and m and m and l. When α is zero, we have a complete symmetry.

Good x is produced by a single profit-maximizing firm in each country with the following cost function:[6]

[6]The numeraire good is produced by perfectly competitive firms operating under constant returns to scale and is freely traded internationally to prevent potential balance of trade issues.

$$C_i(q_i) = cq_i - \delta q_i^2 \text{ where } \left\{ \begin{array}{c} \delta > 0 \text{ if there is Economies of Scale} \\ \delta = 0 \text{ if there is Constant Returns to Scale} \\ \delta < 0 \text{ if there is Diseconomies of Scale} \end{array} \right\}$$

where q_i denotes the total production of good x by firm i: $q_i = \sum_{z=i,j,k} x_{iz}$. Note from the cost function above that δ characterizes the extent of economies of scale. Most of the existing literature employing oligopoly model of trade assumes constant marginal cost and thus assume $\delta = 0$, implying constant returns to scale.[7].

We assume that firms engage in Cournot competition in all markets. In the absence of any FTAs, firm i faces a specific tariff t_{ji} when exporting to country i.[8] Firm i solves the following problem:

$$\max_{x_{ii},x_{ij},x_{ik}} \pi_i = \sum_{z=i,j,k} p_z x_{iz} - cq_i + \delta q_i^2 - \sum_{z \neq i} t_{zi} x_{iz}$$

Given equilibrium prices and export volumes across all countries, the national welfare of country i, denoted by $w_i(\mathbf{t})$ is defined as the sum of consumer surplus (CS_i), producer surplus (PS_i), and tariff revenue (TR_i):

$$w_i(\mathbf{t}) = \underbrace{u_i(x_i(\mathbf{t})) - p_i(\mathbf{t})x_i(\mathbf{t})}_{CS_i(\mathbf{t})} + \underbrace{\pi_i}_{PS_i(\mathbf{t})} + \underbrace{\sum_{z \neq i} t_{iz} x_{zi}}_{TR_i(\mathbf{t})}$$

where $\mathbf{t}$ represents the tariff vector that involves all external tariffs: $\mathbf{t} = (t_{sm}, t_{sl}, t_{ml}, t_{ms}, t_{ls}, t_{lm})$, which represents the tariffs that the three countries impose on each other's exports. Under the No Agreement scenario, given the MFN clause, each country imposes the same tariff on the other two, resulting in the tariff vector $\mathbf{t}^{MFN} = (t_s, t_s, t_m, t_m, t_l, t_l)$. In the case of a Free Trade Agreement between two countries, internal tariffs are eliminated, while each country retains autonomy in setting external tariffs on the non-member country. The non-member country, in turn, imposes the same tariff on both FTA members. For instance, if countries m and ℓ agree to form an FTA, the tariff vector will be $\mathbf{t}^{(m\ell)} = (t_s, t_s, t_{ms}, 0, t_{ls}, 0)$. In the case of a hub and

[7]We employ this cost function for simplicity and more tractable analytical results. However, one should note that this cost function yields equivalent results to those obtained from a general Cobb-Douglas production function. We show that there is a one-to-one correspondence between δ in our cost function and the exponent parameters in the general form of the Cobb-Douglas production function.

[8]It is obvious that $t_{ii} = 0$ since countries do not impose taxes on their local firms and there are no tariffs on the numeraire good (that may be traded internationally in order to balance trade).

spoke agreement, one country forms an FTA with the other two, without the latter two forming an FTA with each other. For example, if country ℓ serves as the hub, the tariff vector would be $\mathbf{t}^{(\ell h)} = (t_{sm}, 0, t_{ms}, 0, 0, 0)$. Lastly, if all countries engage in global free trade, the tariff vector is $\mathbf{t}^{(F)} = (0, 0, 0, 0, 0, 0)$.

Let us define the welfare functions such that, given equilibrium prices and equilibrium export volumes across all countries, the national welfare of country i, denoted by $W_i(\mathbf{t})$, is the sum of consumer surplus, producer surplus, and tariff revenue:

$$W_i(\mathbf{t}) = CS_i(\mathbf{t}) + PS_i(\mathbf{t}) + TR_i(\mathbf{t}) \tag{1}$$

where Consumer Surplus is given by:

$$CS_i(\mathbf{t}) = u(x_i(\mathbf{t})) - p_i(\mathbf{t})x_i(\mathbf{t}) \tag{2}$$

The profit of country i's firm[9] in domestic and foreign markets at given equilibrium prices and export volume. Additionally, the Tariff Revenue of country i is the sum of tariffs imposed on imports from the other two countries to country i, given by:

$$TR_i(\mathbf{t}) = t_{ij}q_{ji}(\mathbf{t}) + t_{ik}q_{ki}(\mathbf{t}) \tag{3}$$

In our three-stage game, the first stage involves the formation of trade agreements. At this stage, considering market size asymmetry (or symmetry) and economies of scale, countries make endogenous decisions on FTA formation. In the second stage, based on the agreements made in the first stage, countries select optimal tariffs with the goal of maximizing welfare. In the final stage, consumption and production take place, guided by the optimal tariffs determined in the previous stage. We employ backward induction to find the subgame perfect Nash equilibrium. The next section explains the strategies employed in our three-stage game to achieve an endogenous equilibrium.

3 Strategies and Equilibrium Concept

In this section, we employ the approach pioneered by Saggi and Yildiz (2010) in applying our three-stage game to the domain of trade liberalization. Under this framework, each country has

[9] $PS_i(\mathbf{t}) \equiv \pi_i(\mathbf{t})$

the freedom to pursue one of three distinct options: (a) no trade liberalization, (b) bilateral trade liberalization, or (c) multilateral trade liberalization. We proceed to analyze how endogenous agreements, influenced by different trade regimes, correspond to strategy sets. Moreover, we introduce a mathematical framework for computing the optimal tariffs under these varying regimes.[10]

3.1 Endogenous Agreements under Bilateralism:

Considering the game of bilateralism, the first phase of the model under analysis is centered on the formation of trade agreements. In this critical stage, countries concurrently make announcements naming their preferred partners for establishing agreements, without committing to any specific trade agreement. It's essential to highlight that announcements are central to this endogenous process of agreement formation. Specifically, within the context of FTAs, a country announces the names of partners with whom it has an interest in forming an FTA. An agreement is formed when there is a concurrence in announcements between countries. For simplicity in this section, we shall use the names i, j, k for countries instead of s, m, l respectively. However, we shall revert to our original notations for the purposes of analysis. Suppose country i makes an announcement, which is denoted by σ_i. Its strategy set of all possible announcements is represented by:

$$\Omega_i = \{\{\phi, \phi\}, \{\phi, j\}, \{\phi, k\}, \{j, k\}\}.$$

Where $\sigma_i = \{\phi, \phi\}$ denotes that country i doesn't want to form an FTA with any country, $\sigma_i = \{j, \phi\}$ and $\sigma_i = \{\phi, k\}$ indicate a desire to form an FTA with country j or k, respectively, and $\sigma_i = \{j, k\}$ denotes an intention to form FTAs with both j and k. This process of matching announcements creates the Endogenous Agreements under Bilateralism.

The FTA game can yield the following trade regimes:

(i) Free trade (F): This arises if and only if $\sigma_i = \{j, k\}$, $\sigma_j = \{i, k\}$, and $\sigma_k = \{i, j\}$. In this case, every country announces the names of both other countries, showing a mutual interest in forming FTAs with one another.

(ii) Bilateral FTA (ij): This occurs if and only if $\sigma_i = \{j, \phi\}$, $\sigma_j = \{i, \phi\}$, and $\sigma_k = \{\phi, \phi\}$.

[10] Stoyanov and Yildiz (2015) and Chi and Yildiz (2022) employed the same approach as ours, but within the context of the political economy literature. Additionally, Missios et al. (2016) utilized the three-stage game approach for comparing customs unions (CUs) and free trade agreements (FTAs), highlighting that CUs prevent global free trade from emerging as a coalition-proof Nash equilibrium, in contrast to FTAs.

This reflects a situation where only countries i and j announce each other's names, indicating mutual interest in forming an FTA exclusively with each other.

(iii) Hub-and-Spoke (ih): This regime arises if $\sigma_i = \{j, k\}$, $\sigma_j = \{i, \phi\}$, and $\sigma_k = \{i, \phi\}$. Here, country i as the 'hub' announces both countries j and k, which act as 'spokes', while countries j and k only announce country i. This leads to country i forming FTAs with both j and k, but countries j and k don't have an FTA with each other as they don't announce each other's names.

(iv) No Agreement ϕ: This is the outcome if $\sigma_i = \{\phi, \phi\}$, $\sigma_j = \{\phi, \phi\}$, and $\sigma_k = \{\phi, \phi\}$. It signifies that there are no matching announcements between any countries, and thus no FTAs are formed.

It is important to further elucidate the mechanics of the announcement game. In this game, countries announce desired partners only, not in favor of specific trade agreements. For instance, in the case of a Bilateral FTA (ij), it arises if the announcements of countries i and j match, regardless of country k's offer. For example, even if country k announces both countries i and j ($\sigma_k = \{i, j\}$), but both i and j do not announce country k ($\sigma_i = \{j, \phi\}$, $\sigma_j = \{i, \phi\}$), the announcements of country k do not match with those of countries i and j. Consequently, only FTA (ij) arises. In this way, there are multiple matching announcement strategy sets for which Bilateral FTA (ij) arises.[11] Similarly, various matching announcement sets can create different types of agreements such as No agreement and Hub-and-Spoke, with the exception of Free Trade, which requires a unique matching announcement set.

In the previous example, a bilateral FTA is in place between countries i and j. However, when country k announces its openness for trade with both countries, this creates an opportunity for countries i and j to initially deviate from the bilateral FTA(ij) with the intent to become hub countries. In the next step of our analysis, we examine the possibilities for countries i and j to engage in deviations in pursuit of becoming hub countries. Notably, in this particular scenario, coalitional deviations are viable for countries aspiring to become hub countries. However, in more general settings, both coalitional and unilateral deviations are conceivable.[12] At this point, a question arises: is this initial deviation self-enforcing, or is further deviation from this initial deviation possible? To address this, we first analyze the Nash equilibria and then

[11] For instance, Bilateral FTA (ij) can arise from other matching announcement strategy sets such as: $\sigma_i = \{j, \phi\}$, $\sigma_j = \{i, \phi\}$, $\sigma_k = \{i, \phi\}$ or $\sigma_i = \{j, \phi\}$, $\sigma_j = \{i, \phi\}$, $\sigma_k = \{\phi, j\}$ or $\sigma_i = \{j, k\}$, $\sigma_j = \{i, \phi\}$, $\sigma_k = \{\phi, \phi\}$, and so on.

[12] Transfers between different coalitions are not allowed in this paper.

allow countries to undertake coalitional deviations in order to isolate the set of coalition-proof Nash equilibria (CPNE), drawing on the approach of Bernheim et al. (1987). They highlight that in many non-cooperative environments, players may discuss strategies but are unable to make binding commitments. Under these circumstances, any significant agreement between the players must be self-enforcing. They further assert that while the Nash best-response property is a necessary condition for self-enforceability, it is not sufficient. They emphasize that it is generally possible for coalitions to orchestrate plausible and mutually beneficial deviations from Nash agreements. This insight is critical, as it suggests that even if an agreement satisfies the Nash equilibrium criteria, it may not necessarily be stable if a coalition has the opportunity to deviate in a manner that benefits all of its members. Consequently, it is essential to consider the self-enforcing nature of the deviations by countries i and j, and assess the potential for further deviations that could destabilize the original agreement.

To understand the concept of self-enforcing agreements, let's continue with our example. In this study, we assess the welfare levels of countries to determine the value of various agreements for them. In our example, when country k announces its openness to trade with both countries i and j, an opportunity arises for both countries to become hub countries, which could enhance their welfare. Imagine a situation where country i forms a coalition with country k to establish a hub-and-spoke agreement, with country i as the hub, taking country j's announcement as given, i.e., $\sigma_j = \{i, \phi\}$. This initial deviation is self-enforcing, as no subset of member countries (i or k) has the option of further deviation, assuming that all countries are worse off under No-Agreement (ϕ), and taking j's announcement as given. This type of initial deviation toward (ih) is self-enforced and acts as a stumbling block in the path toward achieving global free trade. Let's consider a counter-example to further illustrate this point. Imagine three countries are under a free trade agreement. Suppose countries i and j have incentives to collectively deviate from the free trade (F) announcement profile, moving towards a bilateral free trade agreement FTA (ij), while taking country k's announcement as given, i.e., $\sigma_k = \{i, j\}$. The incentives for deviation can be represented as $W_i(ij) = W_j(ij) > W(F)$. The question here again is, is this deviation self-enforcing? The answer is no. Given that $W_i(ih) > W_i(ij) = W_j(ij)$, when taking country k's announcement as given, country i has an incentive to deviate further. This illustrates that the deviation is not self-enforcing as country i is motivated to pursue additional deviations in search of better outcomes. Thus, we can say that the initial deviation to (ij) is not

threatening to Free Trade(F). For the Coalition-Proof Nash Equilibrium (CPNE),[13] we need to examine all initial deviations to ensure that none are self-enforcing. Through this analysis, we can identify stable trade agreements that are immune to self-enforcing coalitional deviations.

3.2 Optimal Tariffs:

Next, we move to stage 2, where the optimal tariff sets are to be determined endogenously. We will solve the mathematical model for each trade regime to determine the endogenous optimal tariffs. In order to guarantee positive exports and tariffs, we impose restrictions on the parameters: $\underline{\delta} < \delta < \bar{\delta}$ and $0 \leq \alpha < \bar{\alpha}$.

Assume that country i imposes a single non-discriminatory (MFN) tariff (in accordance with GATT Article I) $t_i(\phi) = t_{ij}(\phi) = t_{ik}(\phi)$ to maximize its welfare:

$$t_i(\phi) \equiv \arg\max W_i(\phi) \equiv CS_i(\boldsymbol{t}) + PS_i(\boldsymbol{t}) + t_i \sum_{z \neq i} q_{zi}(\boldsymbol{t}) \tag{4}$$

When countries i and j form a bilateral FTA, $t_i(ij) = t_j(ij) = 0$, and each country imposes its individually optimal tariff on country k. Similarly, under the Hub-and-Spoke regime, when country i acts as the hub and countries j and k are the spokes, $t_i(ij)$, $t_j(ij)$, $t_i(ik)$, and $t_k(ik)$ are all set to 0. Spoke countries j and k impose tariffs on each other. We find that the higher the degree of economies of scale, the higher the optimal external tariff.

Upon the formation of bilateral $FTA(ij)$ (in accordance with GATT Article XXIV), member countries impose zero tariffs on each other and impose their individually optimal external tariffs on the non-member.

Under a bilateral $FTA(ij)$, the members, denoted as $v = i, j$, solve:

$$\max_{t_{vk}} \; W_v(ij) \text{ subject to } (t_{ij} = 0, t_{ji} = 0)$$

Under a bilateral $FTA(ij)$, the non-member k solves:

$$\max_{t_{ki}, t_{kj}} \; W_k(ij) \text{ subject to } (t_{ij} = 0, t_{ji} = 0)$$

Likewise, we can find optimal tariffs under all bilateral and Hub-and-Spoke agreements.

Indeed, a pertinent question arises in the formation of bilateral Preferential Trade Agree-

[13] Also called stable equilibria (Saggi and Yildiz, 2006).

ments (PTAs): Do member countries have an incentive to lower their tariffs on non-member countries, and if so, what is the impact of the degree of economies of scale on this? In the relevant literature, the term "Tariff Complementarity Effect" is used to encapsulate this phenomenon. This effect suggests that the establishment of a bilateral $FTA(ij)$ provides each member with a motivation to lower its tariff on the non-member country, in comparison to a situation without any trade agreement.[14] Under this effect, we can observe that:

$$t_{ik}(ij) < t_{ik}(\phi) \quad \& \quad t_{jk}(ij) < t_{jk}(\phi) \tag{5}$$

These inequalities hold in our study, thereby affirming the existence of the Tariff Complementarity Effect in this context. This observation suggests that the formation of an FTA incentivizes member countries to lower their tariffs towards non-member countries. Having established this, we are then prompted to explore how the degree of tariff complementarity changes with the degree of economies of scale. Our observations suggest that tariff complementarity is upheld across all degrees of economies of scale, regardless of whether they are high or low. In fact, the Tariff Complementarity Effect deepens as the degree of economies of scale strengthens. A higher degree of tariff complementarity translates to less discrimination, and vice versa. Intuitively, when FTA member countries become more efficient than non-members, resulting in increased export volumes between them, the exports from non-members to member countries decrease relatively more. This further diminishes the protection incentives of the member countries. The resulting increase in export volumes among member countries leads to trade diversion away from non-member countries. Given this shift, it becomes interesting to analyze the behavior of non-member tariffs. This deeper tariff complementarity effect, as the degree of economies of scale strengthens, can be illustrated as follows: in Figure 1, the horizontal axis represents the degree of economies of scale (δ), and the vertical axis depicts the tariff complementarity effect, i.e., $t_{ik}(\phi) - t_{ik}(ij)$. It is important to explain that due to market size asymmetry, we examined the degree of economies of scale in all bilateral agreements, but the outcome is consistent across all.

[14]The term "Tariff Complementarity Effect" was first introduced by Bagwell and Staiger (1997). Various studies have subsequently supported this theory, including works by Saggi and Yildiz (2010) and Ornelas (2005, 2007). Empirical validation of the Tariff Complementarity Effect has been established by studies such as those conducted by Estevadeordal et al. (2008), Calvo-Pardo et al. (2009), and Bohara et al. (2004). However, this theory has been challenged by researchers such as Limão (2006) and Stoyanov and Yildiz (2015). For further insights into this topic, the reader may refer to Bagwell and Staiger (1997, 1999), Saggi and Yildiz (2009), and Yanase and Tsubuku (2022).

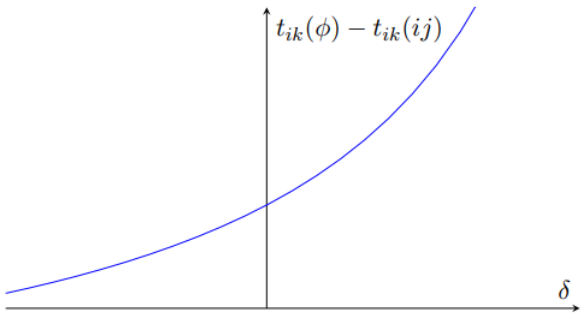

Figure-I: Tariff complementarity effect under FTA

In contrast to member countries, non-members might adjust tariffs depending on the degree of economies of scale. Our observations suggest that an FTA non-member's tariff imposition on an FTA member depends on δ. The specific scenarios are as follows: In contrast to member countries, non-members might adjust tariffs depending on the degree of economies of scale. Our observations suggest that an FTA non-member's tariff imposition on an FTA member depends on δ. The specific scenarios are as follows:

(i) $t_{ki}(\phi) < t_{ki}(ij), \quad$ if $\delta > 0$

(ii) $t_{ki}(\phi) = t_{ki}(ij), \quad$ if $\delta = 0$

(iii) $t_{ki}(\phi) > t_{ki}(ij), \quad$ if $\delta < 0$

When $\delta > 0$, indicating IRS industries , the non-member country (k) has a lower protection incentive and imposes a higher tariff on country i (member country) compared to the MFN tariff. On the other hand, if $\delta < 0$, signifying DRS indsutries, FTA members become more efficient than non-member countries. Consequently, non-members with higher protection incentives lower their tariffs on country i than under the MFN rate. When $\delta = 0$, representing CRS industries, the non-member country (k) is indifferent in its tariff choice compared to the MFN rate. This is also how country k treats another member country j. Thus, a bilateral FTA encourages a member country to lower its tariffs on non-members across all degrees of economies of scale. Meanwhile, a non-member adjusts its tariffs on FTA members based on the nature of economies or diseconomies of scale in the industries. Figure 2 further illustrates the tariff behavior of non-member countries in response to varying degrees of economies of scale. The horizontal axis denotes the degree of economies of scale (δ), while the vertical axis shows the tariff behavior, represented as $t_{ki}(\phi) - t_{ki}(ij)$. Despite market size asymmetry, we explored the

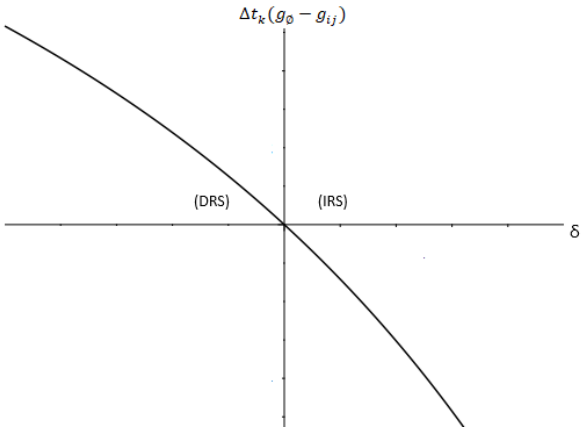

Figure-2: : Nonmember's tariff on member countries

degree of economies of scale across all bilateral agreements and observed consistent results.

Next, we move to the discussion of two countries forming a gated multilateral agreement. If countries i and j agree to sign the gated multilateral agreement (ij^m), they jointly choose their optimal tariffs subject to the constraint that they must not discriminate against country k. That is, in accordance with the MFN clause of the WTO, any tariff cut undertaken by either country i or j must apply equally to imports from both its multilateral member and non-member.

Under a *gated multilateral agreement* (ij^m), the members, denoted as $v = i, j$, jointly solve:

$$\max_{t_v} \left[W_i(ij^m) + W_j(ij^m) \right]$$

Under a *gated multilateral agreement* (ij^m), the non-member k solves:

$$\max_{t_k} W_k(ij^m)$$

Likewise bilateralism, under gated multilateralism (ij^m), each member country lowers its MFN tariff on both member and non-member countries. In this agreement, we observe that:

$$t_v(ij^m) < t_v(\phi), \quad \text{where } v = i, j$$

Our observations suggest that this reduction in MFN tariffs under this agreement holds for all degrees of economies of scale. However, the dependence of both members on the degree of economies of scale for the reduction in MFN tariffs varies and depends on their market

size. For the relatively smaller country among the members, this reduction in external MFN tariffs deepens as the degree of economies of scale strengthens. Suppose i represents the smaller country in the agreement, while j represents the larger country. Then: $\frac{\partial(t_i(\phi)-t_i(ij^m))}{\partial\delta} > 0$. Intuitively, higher economies of scale increase the incentive to lower more the MFN tariffs for the smaller country, as the associated gains from trade increase relative to the losses from foreign competition. This further diminishes the protection incentives of the member countries. The resulting increase in export volumes among member countries leads to trade diversion away from non-member countries. This effect is the same as the deeper tariff complementarity effect observed under bilateralism, as illustrated in Figure 1. However, for the country with the larger market size, the reduction in external MFN tariffs weakens with higher degrees of economies of scale: $\frac{\partial(t_j(\phi)-t_j(ij^m))}{\partial\delta} < 0$. Intuitively, higher economies of scale amplify the internal benefits of the agreement for the larger country, making internal market access more valuable and reducing the incentives to lower external MFN tariffs. The analysis for non-members under gated multilateralism follows the same findings as under bilateralism. As illustrated in Figure 2, non-members adjust their tariffs depending on the degree of economies of scale. Specifically, non-members' tariff impositions on members decrease as economies of scale increase, reinforcing the broader trade diversion effect observed in bilateral agreements.

Finally, we consider the case of two countries forming an *open multilateral agreement*. Under this agreement, countries i and j form a multilateral agreement (ij^{m^o}) and commit to extend any tariff cuts applied to each other to all countries, including the non-member country k.

Under an open multilateral agreement (ij^{m^o}), the members use the common external tariffs, denoted as $t_i = t_j = t$, jointly solve:

$$\max_t \left[W_i(ij^{m^o}) + W_j(ij^{m^o})\right]$$

The non-member k maximizes its own welfare:

$$\max_{t_k} W_k(ij^{m^o})$$

In this scenario, our findings for the implication of economies of scale on external tariffs are consistent with those under multilateralism.

3.3 Welfare effects of FTAs under Bilaterism:

Let's denote country i's welfare as a function of trade agreement r by $w_i(r)$ and let $\Delta w_i(r-v)$ represent the difference between country i's welfare under trade regimes r and v: $\Delta w_i(r-v) \equiv w_i(r) - w_i(v)$. Let $a_i(r-v)$ represent the critical threshold of market size asymmetry at which country i is indifferent between regimes r and v. From now onward, we shall use our original terms "s, m, or l" for countries, instead of "i, j, k", respectively. To obtain the CPNE agreements, we compare the welfare of each country under all its possible announcements. We find that, in forming the CPNE agreements, the degree of economies of scale and a country's market size play decisive roles. In particular, we found that across all bilateral agreements, market size asymmetry plays a crucial role in FTA formation at all degrees of economies of scale. However, this influence is consistent across different levels of economies of scale, but the critical threshold of market size asymmetry varies depending on the level of economies of scale.

Lemma 1: The welfare impact of various types of trade agreements on a small country (s) is as follows:

(i) $0 < \Delta w_s(\phi - v)$, where $v \in \{sm, sl\}$, for all α, for all δ.

(ii) $0 < \Delta w_s(v - F)$, where $v \in \{sm, sl, mh, lh\}$, for all α, for all δ.[15]

(iii) Country s has no incentive to break up any agreement unless it is a spoke.

Part (i) of the lemma suggests that the formation of a bilateral FTA consistently results in a welfare improvement for country s, irrespective of the economies of scale. This indicates that access to larger markets through bilateral FTAs is beneficial to country s under all circumstances. Meanwhile, part (ii) demonstrates the appeal of free trade for the smallest country s, such that there is no incentive to deviate, either unilaterally or jointly, from its free trade announcement. This consistent preference for free trade can be intuitively understood by considering the firm within country s: it finds access to larger markets persistently appealing, unaffected by the degrees of economies of scale. Furthermore, part (iii) reinforces the significance of larger markets to the smallest country, as observed previously in part (ii). Given the larger market sizes of countries l and m compared to country s, any announcement favoring a deviation from any agreement would lead to reduced opportunities for country s to benefit

[15] However, the smallest country (s) consistently prefers to be the hub in comparison to engaging in free trade: $0 < \Delta w_s(F - sh)$, for all α, for all δ.

from these larger markets. Hence, country s shows no incentive to break up any agreement, regardless of the prevailing economies of scale.

Let's now turn our attention to large and medium countries. We've discovered that, given the economies of scale, the size of their markets plays a vital role in the formation of agreements. This fact forms the underlying basis of our subsequent analysis, as defined in Lemma 2 below.

Lemma 2: Given a degree of economies of scale, the welfare impact of various types of trade agreements on countries m and l is as follows:

(i) Countries m and l have incentives to break up an agreement if and only if they are sufficiently large relative to their smaller partners.

(ii) Even though there exist exclusion incentives for countries m and l, these incentives go unexercised due to the flexibility of FTA formation.

(iii) The unilateral incentive of country l to deviate from free trade determines the CPNE conditions of Free Trade. In other words, the free-riding incentive matters.

The intuition behind part (i) is that when the market size (α) and the degree of economies of scale (δ) are high, preferential access to each other's markets becomes a significant issue. In this scenario, the larger country l is giving the smaller country m access to its extensive market in exchange for access to the smaller market of country m relative to l. As such, the incentive for unilateral deviation is quite high for the larger country l, given that it might perceive an imbalance or insufficient return on the agreement.[16] Next, we explore the concepts of exclusion[17] and free-riding effects, as highlighted in part (ii) and part (iii). These effects play a pivotal role in our analysis because they can potentially transform FTAs, or Free Trade Agreements, into stumbling blocks on the path toward achieving global free trade.

We discovered that for significantly high market size α across all degrees of economies of scale δ, there exist exclusion incentives for countries m and l to exclude country s. Specifically,

$$\begin{cases} 0 < \Delta W_m(m\ell - F), \ when\ \alpha_m > \alpha_m(m\ell - F), and \\ 0 < \Delta W_l(m\ell - F) > 0, \ when\ \alpha_\ell > \alpha_\ell(s\ell - F). \end{cases} \tag{6}$$

This suggests that countries m and l prefer the bilateral FTA denoted as (ml) over the

[16] For each agreement involving countries m and l, critical thresholds of market size exist, beyond which these countries will dissolve their agreements.

[17] Exclusion incentive: A pair of countries benefit from jointly deviating from free trade to a bilateral FTA: $\Delta w_v(F - ij) < 0$, where $v = i, j$.

Free Trade Agreement (F), i.e., they choose to exclude country s. Simultaneously, Lemma 1(ii) demonstrates that, across all market sizes α and degrees of economies of scale δ, country s consistently favors participating in the Free Trade Agreement. This preference holds over the option of being excluded from the agreement involving countries (ml). However, the flexibility of FTAs[18] allows countries l and m to jointly announce their deviation from a bilateral FTA with country s, in order to transition into becoming hub countries:

$$\begin{cases} 0 < \Delta w_m(ml - mh), \ when \ \alpha_m < \alpha_m(m\ell - mh) \\ 0 < \Delta w_l(ml - lh), \ when \ \alpha_\ell < \alpha_\ell(m\ell - lh) \end{cases} \tag{7}$$

Also, for the sake of achieving higher welfare, it is more advantageous for country s to position itself as a 'spoke' in the trade network, rather than becoming an outsider from the bilateral FTA (ij).

$$\begin{cases} 0 < \Delta w_s(ml - mh), when \ \alpha_s > \alpha_s(m\ell - mh) \\ 0 < \Delta w_s(ml - lh), when \ \alpha_s > \alpha_s(m\ell - lh) \end{cases} \tag{8}$$

Hence, due to the flexibility of FTA formation, exclusion incentives for countries m and ℓ exist, but they go unexercised. In essence, when compared to free trade, each member country of an FTA possesses a lower ability to manipulate its terms of trade. This concept is precisely encapsulated in Lemma 2 (ii).

To analyze the free-riding effects,[19] we draw upon Lemma 2(i), which illustrates the crucial role that country ℓ, due to its relative size and influence, plays in the formation of global free trade. This influence is particularly evident when considering the condition:

$$0 < \Delta W_\ell(sm - F) \quad \text{when} \quad \alpha_\ell(sm - F) < \alpha_\ell. \tag{9}$$

This shows that, given economies of scale, for very large market sizes, country ℓ doesn't need to provide any preferential access to its market to any country. Instead, it can choose to remain outside any bilateral agreement, emphasizing the importance of the free-riding incentive. In essence, the unilateral incentive of country ℓ to deviate from free trade can influence the CPNE conditions of Free Trade. This suggests that the strategic decisions of country ℓ, with

[18] Flexibility of FTAs: while an FTA member is free to sign another FTA with an existing non-member without needing the consent of an existing FTA partner.

[19] Free-riding incentive: This refers to the benefit a country derives when it deviates from free trade to remain an outsider in a bilateral FTA: $\Delta w_k(F - ij) < 0$.

its substantial market size and influence, can establish the conditions for a stable global trade equilibrium where no country would gain from unilaterally altering its trade policy.

3.4 Nash Equilibria: Bilateralism

Now, we provide a discussion on the Nash equilibria. It is easy to understand that no agreement $\langle\phi\rangle$ is a Nash equilibrium since no country has a unilateral incentive to announce another country's name if it does not receive an announcement in return as a favor. Next, to determine whether each bilateral FTA serves as a Nash equilibrium, we refer to Lemma 1(i). Regardless of any degree of economies of scale, country s does not possess the incentive to unilaterally deviate from its announcement of supporting its bilateral FTA in favor of no agreement. In contrast, according to Lemma 2(i), countries m and ℓ have incentives to break up an agreement if and only if they are sufficiently large relative to their smaller partners. Consequently, both countries m and ℓ are prone to deviate unilaterally in favor of no agreement over their existing bilateral announcement. Thus, $\langle ml\rangle$ is a Nash equilibrium when $\alpha_\ell < \alpha_\ell(\phi - ml)$, $\langle sm\rangle$ is a Nash equilibrium when $\alpha_m < \alpha_m(\phi - sm)$, and $\langle sl\rangle$ is a Nash equilibrium when $\alpha_m < \alpha_m(\phi - sl)$. An important assumption here is that a pure strategy Nash equilibrium of the FTA game is supported by parsimonious announcements, i.e., member countries only announce each other's names and non-member countries do not announce any names. This eliminates the chance of any announcement by a member country to become a hub country.

To determine if any Hub-and-Spoke trade regime is a Nash equilibrium, we observe that country ℓ has no incentive to announce its intention to become a spoke country for countries s and m for all α and δ. This is because country ℓ would have to give up more of its market power. Consequently, this eliminates the possibility that $\langle mh\rangle$ or $\langle sh\rangle$ can be Nash equilibria:

$$0 < W_\ell(sm - zh),\ z = s, m,\ \forall a, \forall\delta \tag{10}$$

However, the largest market of country ℓ is attractive to both other countries compared to any announcement of non-membership in any bilateral FTA involving country ℓ. Even with diseconomies of scale, both countries can leverage their access to this largest market according to their profit requirements:

$$\begin{cases} W_s(m\ell - \ell h) < 0, \ when \ \alpha_s > \alpha_s(m\ell - \ell h), \\ W_m(s\ell - \ell h) < 0, \ when \ \alpha_m > \alpha_m(s\ell - \ell h) \end{cases} \tag{11}$$

Here $\alpha_s(m\ell - \ell h) < \alpha_m(s\ell - \ell h)$. On the other side, country ℓ has an incentive to announce itself as a hub country up to some threshold value of a and $\forall\delta$:

$$\begin{cases} W_\ell(m\ell - \ell h) < 0, \ when \alpha_\ell < \alpha_\ell(m\ell - \ell h), \\ W_\ell(s\ell - \ell h) < 0, \forall a, \forall\delta \end{cases} \tag{12}$$

Clearly, we can see that country ℓ announces itself as a hub country rather than remaining in any bilateral agreement for $\alpha_\ell < \alpha_\ell(m\ell - \ell h)$. Since $\alpha_s(m\ell - \ell h) < \alpha_m(s\ell - \ell h)$, this hub regime of country ℓ will be a Nash equilibrium for $\alpha_m > \alpha_m(s\ell - \ell h)$ and $\alpha_\ell < \alpha_\ell(s\ell - \ell h)$. Given economies of scale δ, for very high α, and considering that the market size of country s is very small, country ℓ has no incentive to grant preferential access to its market, and the incentive for it to become a hub is reduced. On the other hand, both countries s and m can decline the role of a spoke only if the deviation in market size is minimal. When this deviation is large, these countries are keen to obtain preferential access to the largest market and opt to become a spoke country.

Lastly, we turn to the Nash equilibrium condition for global free trade. Note that part (ii) of Lemma 1 informs us that country s has no incentive for unilateral deviation from a free trade announcement. Moreover, Lemma 2(iii) and equation (6) indicate that for country ℓ, the free-riding incentive is relevant when $\alpha_\ell > \alpha_\ell(sm - F)$. Hence, country ℓ has an incentive to unilaterally sever its FTA links with both s and m. As a result, no country has a unilateral incentive to deviate from their announcement profiles leading to global free trade for $\alpha_\ell < \alpha_\ell(sm - F)$.

Proposition 1: Given a degree of economies of scale δ, the following announcement profiles are Nash equilibria of the FTA game:

(i) $\Omega^\phi \equiv \{\alpha_s = \{\phi, \phi\}, \ \alpha_m = \{\phi, \phi\}, \ \alpha_\ell = \{\phi, \phi\}\}$ leading to $\{\phi\}$, $\forall a$.

(ii) $\Omega^{ms} \equiv \{\alpha_s = \{m, \phi\}, \ \alpha_m = \{s, \phi\}, \ \alpha_\ell = \{\phi, \phi\}\}$ leading to $\{ms\}$ for $\alpha_m < \alpha_m(\phi - ms)$. Likewise, $\Omega^{m\ell}$ leading to $\{m\ell\}$ for $\alpha_\ell < \alpha_\ell(\phi - m\ell)$. $\Omega^{s\ell}$ leading to $\{s\ell\}$ for $\alpha_\ell < \alpha_\ell(\phi - s\ell)$.[20]

[20] $\Omega^{m\ell} \equiv \{\alpha_s = \{\phi, \phi\}, \ \alpha_m = \{\phi, \ell\}, \ \alpha_\ell = \{\phi, m\}\}$, $\Omega^{s\ell} \equiv \{\alpha_s = \{\phi, \ell\}, \ \alpha_m = \{\phi, \phi\}, \ \alpha_\ell = \{s, \phi\}\}$.

(iii) $\Omega^{\ell h} \equiv \{\alpha_s = \{\phi, \ell\},\ \alpha_m = \{\phi, \ell\},\ \alpha_\ell = \{s, m\}\}$ leading to $\{\ell h\}$, for $\alpha_m > \alpha_m(s\ell - \ell h)$ and $a_l < a_l(s\ell - \ell h)$.

(iv) $\Omega^{F} \equiv \{\alpha_s = \{m, \ell\},\ \alpha_m = \{s, \ell\},\ \alpha_\ell = \{s, m\}\}$ leading to $\{F\}$, for $\alpha_\ell < \alpha_\ell(sm - F)$.

Proposition 3 shows that multiple Nash equilibria exist. To address this multiplicity, we now turn our focus to refining the Nash equilibria that are resistant to self-enforcing coalitional deviations.

3.5 Stable Nash equilibria:

Let us start with the global free trade (F). We allow for joint/unilateral deviation from the announcement profile (F) to one leading to any bilateral FTAs, Hub-and-Spoke, or no agreement. We rule out the possibility of each type of announcement of initial deviation to show that (F) is stable. From Lemma 1(ii), country s has no incentive to deviate from their announcements from (F), and country m always has a stronger incentive under the announcement profile (F) than under no agreement, (ϕ):

$$0 < W_m(\phi - F),\ \forall a,\ \forall \delta \tag{13}$$

Hence, there exists no incentive to deviate unilaterally or jointly to announcement profiles that lead to a deviation from (F) to (ϕ).

To determine whether the announcement profile(s) leading to an initial deviation for Hub-and-Spoke is self-enforcing, we note that (sh) and (mh) are not even Nash equilibria. Therefore, we need only consider the announcement profile(s) leading to an initial deviation from (F) to (ℓh). Lemma 1(ii) suggests that country s always has more incentive under a free trade agreement, but country m has an incentive to initially deviate from the announcement profile of (F) to (ℓh), especially for high α:

$$0 < W_m((\ell h) - (F)),\ \text{when } \alpha_m < \alpha_m((\ell h) - (F)),\ \forall \delta \tag{14}$$

This initial deviation announcement profile for high α is self-enforced as any further deviation of the country toward $(s\ell)$ would lead it into a free-riding announcement profile. Lemma 2(iii) indicates that the free-riding incentive matters only for country ℓ. Hence, we can say that if $\alpha_m < \alpha_m((\ell h) - (F))$, then the initial deviation announcement profile from (F) to (ℓh) is also

not possible.

Finally, we shall check whether any two countries have an incentive to jointly deviate from the announcement profile of global free trade to any bilateral FTAs. According to Lemma 1(ii), country s will not be part of any such initial deviation. Lemma 2(iii) states that the free-riding incentives of country ℓ determine the CPNE condition of global free trade. This free-riding incentive for country ℓ holds for $\alpha_\ell > \alpha_\ell(sm - F)$ and is self-enforced, as country ℓ does not choose to become the spoke of any smaller country. Further, $\alpha_\ell(sm - F) < \alpha_m(\ell h - F)$. Hence, the CPNE condition for global free trade is defined as when $\alpha_\ell < \alpha_\ell(sm - F)$, the announcement profile leading to global free trade is the unique CPNE. The most bottom area in Figure 1 demonstrates that Global FT is the unique CPNE when $\alpha_\ell < \alpha_\ell(sm - F)$.

Next, we examine whether (sm) can also be stable when (F) is out of reach. We need to rule out that the following announcement profiles leading to coalitional or unilateral deviations are self-enforced:

- JD1: Given the strategy of country s, the announcement profiles of countries m and ℓ lead to initial deviation from (sm) to (mh).
- JD2: Given the strategy of country s, the announcement profiles of countries m and ℓ lead to initial deviation from (sm) to $(m\ell)$.
- JD3: Given the strategy of country m, the announcement profiles of countries s and ℓ lead to initial deviation from (sm) to (sh).
- JD4: Given the strategy of country m, the announcement profiles of countries s and ℓ lead to initial deviation from (sm) to $(s\ell)$.
- UD5: Given the strategies of countries s and ℓ, the announcement profile leads to a unilateral deviation of country m from (sm) to (ϕ).
- UD6: Given the strategies of countries m and ℓ, the announcement profile leads to a unilateral deviation of country s from (sm) to (ϕ).

We can immediately rule out joint deviations JD1 and JD3, as (mh) and (sh) are not Nash equilibria. Country ℓ lacks the incentive for joint deviations in JD2 and JD4 for high values of α:

$$\begin{cases} W_\ell(sm - m\ell) < 0, \ when \alpha_\ell < \alpha_\ell(sm - m\ell), and \\ W_\ell(sm - s\ell) < 0, \ when \alpha_\ell < \alpha_\ell(sm - s\ell). \end{cases} \tag{15}$$

However, $\alpha_\ell(sm - m\ell) < \alpha_\ell(sm - F)$ and $\alpha_\ell(sm - s\ell) < \alpha_\ell(sm - F)$ represent the region where (F) is the unique CPNE. Therefore, these two joint deviations can also be ruled out. We can rule out UD6 due to Lemma 1(i), and Lemma 2(i) indicates that UD5 is crucial for determining the CPNE condition for (sm). Country m has an incentive to break up an agreement with country s unilaterally for very high α, when $\alpha_m > \alpha_m(\phi - sm)$. We can conclude that, when (F) is out of reach, (sm) is the CPNE only when the degree of market size asymmetry is within a medium range, specifically when $\alpha_\ell(sm - F) < \alpha_\ell < \alpha_m(\phi - sm)$.

Next, we examine the no-agreement arrangement (ϕ) as a candidate for a stable Nash equilibrium for very high α and corresponding δ. Every bilateral FTA formation offers higher incentives for member countries compared to no agreement. These increased incentives for member countries exist either for $\forall a$ and $\forall \delta$ or up to some threshold values of α and corresponding δ. To demonstrate that no-agreement (ϕ) is CPNE, we need to rule out the possibility that the following initial deviations are self-enforced.

- (Jϕ1): The announcement profile leads to a joint deviation of all three countries from (ϕ) to (ℓh).
- (Jϕ2): The announcement profile leads to a joint deviation of countries s and ℓ from (ϕ) to $(s\ell)$.
- (Jϕ3): The announcement profile leads to a joint deviation of countries m and ℓ from (ϕ) to $(m\ell)$.

(Jϕ1) holds for $\alpha_\ell < \alpha_\ell(\phi - \ell h)$. From equation (8), country ℓ has a higher incentive for further unilateral deviation[21] to $(m\ell)$ from (ℓh) for $\alpha_\ell > \alpha_\ell((m\ell) - (\ell h))$, considering the strategies of countries s and m as given. Hence, (Jϕ1) is not self-enforced when $\alpha_\ell > \alpha_\ell((m\ell) - (\ell h))$. (J$\phi$3) is valid when $\alpha_\ell < \alpha_\ell(\phi - m\ell)$.[22] This is not self-enforced for high α due to Lemma 2(i). (Jϕ1) is no longer significant, as $\alpha_\ell(\phi - s\ell) < \alpha_\ell(\phi - m\ell)$. Hence, we can conclude that no agreement (ϕ) is CPNE for $\alpha_\ell > \alpha_\ell(\phi - m\ell)$ (i.e., market size asymmetry is sufficiently large), and this holds only when $\delta > 0$.

[21] This further deviation can be a joint deviation involving both countries m and ℓ, as $W_m((m\ell) - (\ell h)) > 0$, $\forall a$ and $\forall \delta$.

[22] $W_\ell((\phi) - (m\ell)) < 0$ when $\alpha_\ell < \alpha_\ell(\phi - m\ell)$, and $W_m((\phi) - (m\ell)) < 0$, $\forall a$ and $\forall \delta$.

Lastly, we check whether (ℓh) is a stable Nash equilibrium. The remaining range for α is $\alpha_m(\phi - sm) < \alpha_\ell < \alpha_\ell(\phi - m\ell)$. Outside this range, (F), (sm), or (ϕ) are CPNE for different ranges of α. Therefore, we shall not consider the announcement profiles that lead to these three initial deviations for this range of α. Instead, we need to focus solely on the following two potential announcement profiles that lead to joint deviations from (ℓh):

- (Uℓh1): A unilateral deviation to the announcement profile of country ℓ that leads to a deviation from (ℓh) to $(m\ell)$.
- (Uℓh2): A unilateral deviation to the announcement profile of country ℓ that leads to a deviation from (ℓh) to $(s\ell)$.

(Uℓh1) is self-enforced for $\alpha_\ell < \alpha_\ell(\phi - m\ell)$. Hence, (ℓh) is not a CPNE for $\alpha_m(\phi - sm) < \alpha_\ell < \alpha_\ell(\phi - m\ell)$. Now, we are prepared to conclude the stability of free trade in the following proposition.

- **Proposition 2:**

  (i) Global free tradeis the unique CPNE unless market size asymmetry is sufficiently large ((F)is more likely as the degree of economies of scale rises) when $\alpha_\ell < \alpha_\ell(sm - F)$.

  (ii) When (F) is out of reach, (sm) is CPNE only when the degree of market size asymmetry is at a medium range, i.e., when $\alpha_m(\phi - sm) > \alpha_\ell > \alpha_\ell(sm - F)$.

  (iii) No agreement is CPNE only when there is a large degree of diseconomies of scale and market size asymmetry is sufficiently large.

We will use Figure III to illustrate the intuitions behind Proposition 2. In this figure, the horizontal axis represents economies of scale, denoted as δ, and the vertical axis indicates market size, represented by α. Let's begin by concentrating on the region where the announcement profile leading to global free trade, (F), is a unique CPNE. This region is depicted as the area beneath the lowest IC curve (i.e., $\alpha_\ell < \alpha_\ell(sm - F)$) in Figure III. It reveals the following results: First, For high α, i.e., when country ℓ is significantly larger compared to m and country s is relatively smaller compared to m, and there are IRS, the achievement of global free trade (F) is most likely. In this scenario, the market power of countries is not crucial for any country. The IRS nature of firms provides a higher incentive for each country to grant other countries

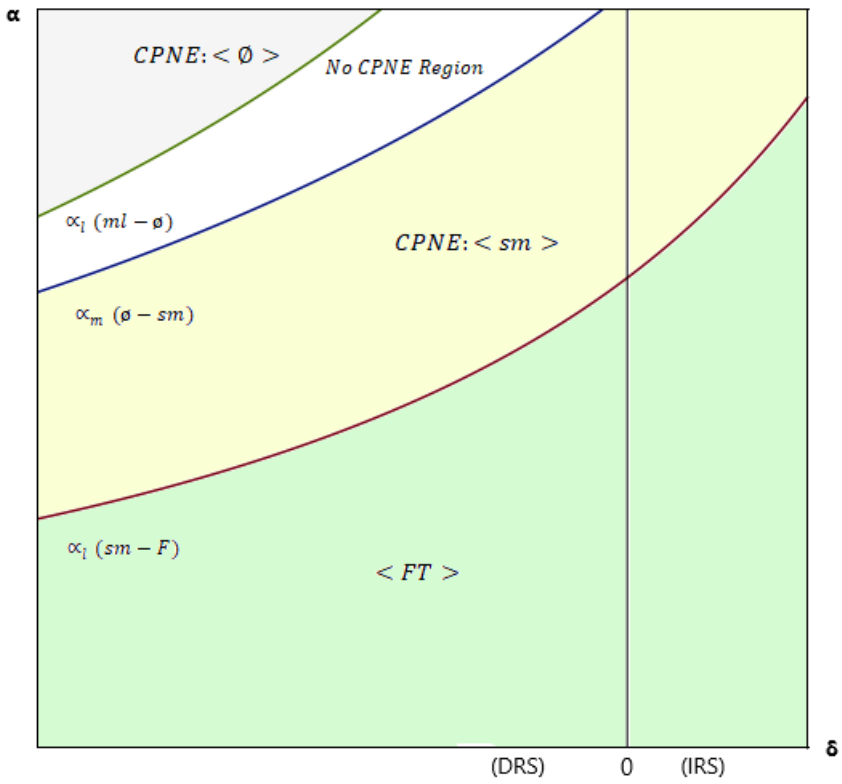

Figure-III: CPNE Bilateralism

access to its market. Second, On the other side of the graph, as the degree of diseconomies rises, making production costs more expensive, the market power of countries gains prominence, and the likelihood of achieving global free trade (F) diminishes. Even for small α values, when a large country allows access to its market for a smaller one, it receives less in return, thereby increasing the large country's incentive for unilateral deviation and heightening the appearance of free-riding incentives.

Next, we discuss when the announcement profile leading to bilateral FTA (sm) arises as the CPNE. This is the area between the second and most bottom ICs (i.e., $\alpha_\ell(sm - F) < \alpha_\ell < \alpha_m(\phi - sm)$). As we discussed above, the free-riding incentive appears in this region, where the large country has an incentive to stay out of any bilateral FTA. First, we focus on the free-riding incentive in IRS/CRS, where α is a symmetric deviation between the market sizes of $\{m, \ell\}$ and $\{m, s\}$. A significantly high α increases country ℓ's market size substantially, while it reduces the market size of country s to a much smaller scale. Essentially, a substantial α amplifies the market power of country ℓ and diminishes that of country s. Producer surplus cannot increase in country ℓ through exports anymore, as it entails conceding its market to firms from the

other two countries, which collectively represent an equal or smaller market size. Consequently, free-riding becomes the optimal strategy for country ℓ in this area. Conversely, in scenarios with a very high degree of diseconomies of scale, even if α is not excessively high, a significant deviation still exists. In this situation, country ℓ is inclined towards free-riding. Participation in (F) cannot bolster producer surplus in country ℓ due to the prevailing Diseconomies of Scale (DRS). Regarding the (sm) is a CPNE following arguments are also significant:

(a) In the graph region corresponding to IRS/CRS, the announcement profile leading to the bilateral FTA (sm) is a CPNE. Although lemma 2(i) provides an avenue for country m to further deviate and potentially dissolve this agreement, the IRS nature of the market prompts country s to maintain the agreement with a smaller market country. This preference stems from the fact that, within an IRS context, very high α values result in significantly small market sizes for individual countries. By enlarging their market size through the bilateral agreement, both countries can enhance their producer surplus.

(b) In contrast, within the DRS zone, lemma 2(i) comes into effect, leading country m to break the agreement when α values are very high. This decision is influenced by the DRS effect, which could negatively impact the producer surplus due to a high degree of δ. Essentially, under significant diseconomies of scale, the potential harm to producer surplus overrides the benefits of maintaining the bilateral agreement.

Now, we discuss the conditions under which the announcement profile leading to no agreement (ϕ) emerges as the CPNE, specifically when $\alpha_\ell > \alpha_\ell(m\ell - \phi)$. This refers to the area above the uppermost IC (on the top left side) in Figure III. Analysis of this area reveals the following results: In the case of IRS/CRS, no agreement (ϕ) does not arise as the CPNE. We have already explored how IRS/CRS dynamics provide incentives for member countries to maintain their bilateral FTA agreements rather than dissolving them. Conversely, under conditions of a very high degree of diseconomies of scale coupled with a high α, both countries m and ℓ seek to protect their producer surplus. They opt not to allow access to each other's markets for exports due to the pronounced DRS effect. This protective stance stems from the high costs and risks associated with production and trade in a high DRS environment, prompting these countries to refrain from entering into agreements that might compromise their economic stability.

Finally, we delve into the specific area delineated by $\alpha_m(\phi - sm) < \alpha_\ell < \alpha_\ell(m\ell - \phi)$, where no distinct Nash equilibrium materializes as the CPNE. In this zone, every announcement profile leading to any form of agreement is superseded by another self-enforced initial deviation.

Notably, this scenario unfolds exclusively under the DRS nature of production. More precisely, it prevails under conditions of very high α even with a low degree of diseconomies of scale, as well as when faced with an extremely high degree of diseconomies coupled with a substantial α value. Above this specific sector, we've identified a state where (ϕ) is the CPNE, while beneath it, we observe the stabilization of (sm) and (F) as enduring Nash equilibria. We can interpret this middle ground as a pressure point where the interplay of market size and the degree of diseconomies of scale can unsettle countries in their foundational agreements, encouraging them to seek incentives (either as participating members or as free riders) within alternative accords. This stress continues to endure, particularly for countries ℓ or m, ultimately resulting in a situation where no consensus agreement emerges as the CPNE amid increasing market sizes and intensifying degrees of diseconomies of scale. Essentially, this point represents a threshold where the pressures from market forces and the complexities of production scales drive countries to a tipping point, after which sustaining the initial or alternative agreements becomes unfeasible, hence, no particular pact surfaces as a definitively stable equilibrium.

Next, we consider a scenario where countries are not able to form PTAs (as Article XXIV doesn't exist). In this case, we move to our Multilateralism game and examine the following question: how are the prospects of attaining global free trade affected if countries cannot form bilateral FTAs?

4 Endogenous MFN- based Agreement Formation:

In this section, we analyze the equilibrium theory in a Multilateral World, where countries can form trade agreements solely on an MFN basis, particularly under conditions where Article-XXIV does not exist, preventing the pursuit of free trade agreements' liberalization. Instead, countries are confined to trade liberalization on a multilateral front, a context we refer to as the 'multilateralism game.' Within this framework, we consider both distinctive approaches to Multilateralism in a three-country setting: open multilateralism and gated multilateralism. This analysis aims to compare the equilibria emerging with and without the capacity to form FTAs to enhance the prospects of global free trade.

4.1 Gated Multilateralism

We will initially explore equilibrium analysis for gated multilateralism, where countries have the authority to exclude others from their agreements. Under a multilateral approach to trade liberalization, Country s makes an announcement, denoted by α_s. Its strategy set, α_s, is defined as follows:

$$\alpha_s = \{\{\phi, \phi\}, \{M_m, \phi\}, \{\phi, M_\ell\}, \{M_m, M_\ell\}\}.$$

Where $\alpha_s = \{\phi, \phi\}$ indicates that country s does not want to form a trade agreement with any country. $\alpha_s = \{M_m, \phi\}$ denotes that country s forms a trade agreement with country m on an MFN basis. $\alpha_s = \{\phi, M_\ell\}$ signifies that Country s forms a trade agreement with country ℓ on an MFN basis. $\alpha_s = \{M_m, M_\ell\}$ indicates that Country s wishes to form a trade agreement with all parties on an MFN basis. Similarly, we can define strategy sets α_m and α_ℓ for countries m and ℓ, respectively.

The multilateral game can yield the following trade regimes:

(i) F^m arises if and only if: $\alpha_s = \{M_m, M_\ell\}$, $\alpha_m = \{M_s, M_\ell\}$, and $\alpha_\ell = \{M_s, M_m\}$, meaning every country announces all others' names.

(ii) sm^m arises if and only if: $\alpha_s = \{M_m, \phi\}$, $\alpha_m = \{M_s, \phi\}$, and $\alpha_\ell = \{\phi, \phi\}$, indicating there is only one pair, s and m, of matching announcements. Similarly, we can define $s\ell^m$ and $m\ell^m$. It is important to note that under such agreements, both member countries jointly maximize their payoffs, subject to the constraint that they must extend the same tariff concessions to the external (non-member) country.

(iii) No Agreement ϕ^m arises if: $\alpha_s = \{\phi, \phi\}$, $\alpha_m = \{\phi, \phi\}$, and $\alpha_\ell = \{\phi, \phi\}$. This scenario occurs if, and only if, there is no matching announcement.

We will employ the same approach to finding stable Nash equilibria as we did in the FTA game previously. Notably, there is no hub-and-spoke trade regime in this game, implying fewer initial deviations from any agreement. For example, for the free trade agreement F^m, we permit the joint/unilateral initial deviation from the announcement profile F^m to one leading to any bilateral agreement or no agreement at all.

Subsequently, we present a lemma that will aid in establishing the CPNE within the context of the Gated Multilateralism game:

Lemma 3: Given any degree of economies of scale δ, the following holds:

(i) Country s and m have no incentives to unilaterally or jointly deviate from free trade.

(ii) Larger countries have incentives to break up two-country agreements with a smaller partner if they are sufficiently large to unilaterally or jointly deviate from free trade.

Part *(i)* indicates that for countries s and m, the producers' surplus for firms in both countries increases upon entering the largest market of country ℓ. If α reaches a very high value, making country ℓ significantly larger compared to the other two countries, or if there is a higher degree of diseconomies of scale, the market share for both smaller countries diminishes. Hence, exclusion incentives do not exist. Evidently, the role of country ℓ becomes decisive, and the CPNE condition for free trade relies exclusively on the decisions made by country ℓ. Next, we identify the threshold value that would compel country ℓ to exit the free trade agreement. The incentive for free-riding becomes significant for country ℓ, as it considers restricting access to its larger market for smaller countries in exchange for less:

$$\alpha_\ell(sm^m - F^m) < \alpha_\ell(m\ell^m - F^m) < \alpha_\ell(s\ell^m - F^m) \tag{16}$$

Part *(ii)* demonstrates the market power of larger countries in two-country agreements involving a smaller partner. These larger countries can dissolve such agreements if their market size is sufficiently dominant. The intuition behind market power remains consistent with our earlier discussions. However, in these agreements, while both countries work together to jointly maximize their payoffs, market size remains a crucial factor for the producer surplus of the larger countries.

$$\begin{cases} W_m(sm^m - \phi^m) < 0, \ when \ \alpha_m > \alpha_m(sm^m - \phi^m), \\ W_\ell(s\ell^m - \phi^m) < 0, \ when \alpha_\ell > \alpha_\ell(s\ell^m - \phi^m), \\ W_\ell(m\ell^m - \phi^m) < 0, \ when \alpha_\ell > \alpha_\ell(m\ell^m - \phi^m). \end{cases} \tag{17}$$

The Nash equilibria under the gated multilateralism game are defined in the following proposition.

Proposition 3: Given a degree of economies of scale δ, the following announcement profiles are Nash equilibria of the gated multilateralism game:

(i) $\Sigma^\phi \equiv \{\alpha_s = \{\phi, \phi\}, \alpha_m = \{\phi, \phi\}, \alpha_\ell = \{\phi, \phi\}\}$ leading to $\{\phi^m\}$, $\forall a$.

(ii) $\Sigma^{ms} \equiv \{\alpha_s = \{M_m, \phi\}, \alpha_m = \{M_s, \phi\}, \alpha_\ell = \{\phi, \phi\}\}$ leading to $\{ms^m\}$ for $\alpha_m < \alpha_m(\phi - ms^m)$. Likewise, $\Sigma^{m\ell}$ leading to $\{m\ell^m\}$ for $\alpha_\ell < \alpha_\ell(\phi - m\ell^m)$, and $\Sigma^{s\ell}$ leading to $\{s\ell^m\}$ for $\alpha_\ell < \alpha_\ell(\phi - s\ell^m)$.[23]

(iii) $\Sigma^F \equiv \{\alpha_s = \{M_m, M_\ell\}, \alpha_m = \{M_s, M_\ell\}, \alpha_\ell = \{M_s, M_m\}\}$ leading to $\{F^m\}$, for $0 < a < \alpha_\ell(sm^m - F^m)$.

To address the multiplicity of Nash equilibria presented in Proposition 3, we narrow our focus to refining those equilibria that are immune to self-enforcing coalitional deviations.

4.1.1 Stable Nash equilibria - Multilateralism game:

Let's begin with the announcement profile leading to the free trade agreement (F^m). We allow the following joint or unilateral deviations from the announcement profile (F^m):

- (DMF1): Unilateral or joint deviations in the announcement profile leading to any of the following bilateral agreements: $(ms^m), (m\ell^m), (s\ell^m)$.
- (DMF2): A joint deviation in the announcement profile leading to no agreement, represented as (ϕ^m).

From Lemma 3*(i)*, we deduce that countries s and m have no incentive to deviate from the announcement profile (F^m). This deduction allows us to dismiss deviations to announcement profiles $(m\ell^m)$ and $(s\ell^m)$ under DMF1 and deviations leading to no agreement (ϕ^m) under DMF2. However, Lemma 3*(ii)* reveals an incentive for country ℓ to deviate unilaterally from the free trade announcement profile to exploit free-riding opportunities, especially when $\alpha_\ell < \alpha_\ell(sm^m - F^m)$. Consequently, country ℓ defines the CPNE condition $\alpha_\ell < \alpha_\ell(sm^m - F^m)$ for the announcement profile that leads to free trade (F^m) in the gated multilateralism game, given δ. The most bottom region is portrayed in Figure-II, illustrating this scenario. It's observable that the IRS region possesses a greater extension of this area compared to the DRS region.

Next, we examine whether a two-country MFN-based agreement can also be stable when (F^m) is out of reach. It is necessary to rule out any unilateral deviations toward (ϕ^m) by any member country, ensuring self-enforcement. Lemma 3*(ii)* shows that country s can't make unilateral deviations toward (ϕ^m). However, countries m and ℓ have incentives to break up their

[23] $\Sigma^{m\ell} \equiv \{\alpha_s = \{\phi, \phi\}, \alpha_m = \{\phi, M_\ell\}, \alpha_\ell = \{\phi, M_m\}\}$, $\Sigma^{s\ell} \equiv \{\alpha_s = \{\phi, M_\ell\}, \alpha_m = \{\phi, \phi\}, \alpha_\ell = \{M_s, \phi\}\}$.

two-country agreements with a smaller partner when they are sufficiently large. It is important to note the following condition:

$$\alpha_\ell(\phi^m - s\ell^m) \equiv min\{\alpha_\ell(\phi^m - s\ell^m), \alpha_\ell(\phi^m - m\ell^m), \alpha_m(\phi^m - sm^m)\} \tag{18}$$

We can establish that within the range $\alpha_\ell(sm^m - F^m) < \alpha_\ell < \alpha_\ell(\phi^m - s\ell^m)$ and for all values of δ, all two-country MFN-based agreements stand as CPNE. This scenario is depicted in the second-to-last strip from the bottom in Figure II, illustrating regions where two-country MFN-based agreements are stable. Notably, in the IRS region with a very low degree of economies of scale, no additional CPNE arises when these agreements are deemed CPNE. The joint welfare of member countries remains elevated with a low degree of economies of scale IRS, where the cost per unit is minimal. In this context, granting access to smaller markets within a relatively larger market provides incentives to the latter country. Consequently, the market power of the larger country diminishes in significance, and there is no compelling incentive for unilateral deviation by any country. However, as this degree increases, the welfare of the larger country starts to decline within the joint welfare configuration with country s, and the incentive for unilateral deviation by country ℓ intensifies. We observe that the area of this strip contracts as the degree of economies of scale escalates. Specifically, under conditions of diseconomies of scale, there is a substantial incentive for the larger country to unilaterally deviate from $(s\ell^m)$ to (ϕ^m).

Now, we shift our focus to identifying the CPNE regions for the announcement profiles leading to $(m\ell^m)$ and (sm^m). Country m initially exhibits a stronger incentive for joint welfare with country s. However, this incentive declines in the face of significantly high "a" or elevated δ levels, inclining it towards a unilateral deviation to (ϕ^m). Similarly, the stability of agreements under $(m\ell^m)$ is compromised due to country ℓ's potential choice of unilateral deviation, particularly when faced with high values of "a" and δ. The CPNE ranges for both $(m\ell^m)$ and (sm^m) are given:

$$\begin{cases} W_m(\phi^m - sm^m) < 0, \ when \alpha_\ell(sm^m - F^m) < \alpha_m < \alpha_m(\phi^m - sm^m), \\ W_\ell(\phi^m - m\ell^m) < 0, \ when \alpha_\ell(sm^m - F^m) < \alpha_\ell < \alpha_\ell(\phi^m - m\ell^m). \end{cases} \tag{19}$$

Referring to Figure II, we observe a notable intersection of both ICs at a specific point, $\delta^{critical}$. This intersection highlights distinct regions: one where only (sm^m) qualifies as CPNE

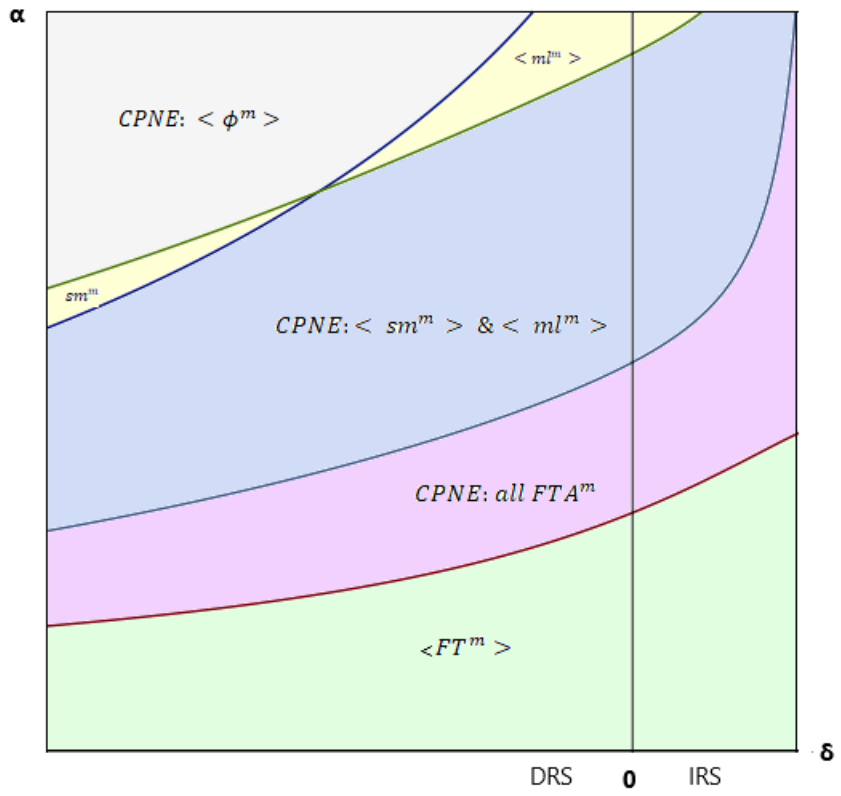

Figure-II: CPNE Multilateralism(Gated)

and another where only $(m\ell^m)$ does. The criteria defining these areas are as follows:

For (sm^m) to be the sole CPNE, the range is specified by:

$$W_m(\phi^m - sm^m) < 0,\ when \alpha_\ell(\phi^m - m\ell^m) < \alpha_m < \alpha_m(\phi^m - sm^m)\ ,\ \delta > \delta^{critical}. \tag{20}$$

Conversely, for $(m\ell^m)$ to be the exclusive CPNE, the conditions are:

$$W_\ell(\phi^m - m\ell^m) < 0,\ when\ \alpha_m(\phi^m - sm^m) < \alpha_\ell < \alpha_\ell(\phi^m - m\ell^m)\ ,\ \delta^{critical} > \delta. \tag{21}$$

The phenomenon occurs because the joint market size of $(m\ell^m)$ is larger than that of (sm^m). As δ increases, particularly in the DRS region, this larger market size leads to more expensive production costs. Consequently, the CPNE area for $(m\ell^m)$ diminishes more rapidly than that for (sm^m) with an increase in δ under DRS conditions for a given α.

Lastly, we examine whether the announcement profile leading to no agreement, (ϕ^m), can emerge as a CPNE. Intuitively, if both the free trade and two-country MFN-based announce-

ment profiles are unattainable, the only feasible CPNE is represented by the 'no agreement' announcement profile. This scenario is depicted in Figure II, observable in the top right corner. Mirroring our arguments from the bilateral FTA game, we assert that (ϕ^m) as a CPNE is plausible in situations of diseconomies of scale, especially for significantly high values of "a" and elevated δ levels. The range where (ϕ^m) emerges as CPNE is defined by:

$$\{\alpha_\ell \mid \alpha_\ell > \alpha_\ell(\phi^m - m\ell^m), \text{and}, \delta < \delta^{critical}\}, \text{and}, \{\alpha_m \mid \alpha_m > \alpha_m(\phi^m - sm^m), \text{and}, \delta > \delta^{critical}\} \tag{22}$$

Now, we conclude the stability of the equilibrium trade regimes for our game of multilateralism in the following proposition.

Proposition 4: Given the degree of economies of scale, δ, the following holds:

1. *(i)* Global free trade is the unique CPNE unless market size asymmetry is large (i.e., $< F^m >$ is more likely as the degree of economies of scale rises) when $\alpha_\ell < \alpha_\ell(sm^m - F^m)$.

2. *(ii)* When $< F^m >$ is out of reach, all possible two-country MFN-based agreements are CPNE only when the degree of market size asymmetry is in a low to medium range, specifically when $\alpha_\ell(\phi^m - s\ell^m) > \alpha_\ell > \alpha_\ell(sm^m - F^m)$.

3. *(iii)* When $< F^m >$ and $< s\ell^m >$ are out of reach, $< sm^m >$ and $< m\ell^m >$ are CPNE only when the degree of market size asymmetry is in the medium range. Specifically, $\alpha_m(\phi^m - sm^m) > \alpha_m > \alpha_\ell(\phi^m - s\ell^m)$ and $\alpha_\ell(\phi^m - m\ell^m) > \alpha_\ell > \alpha_\ell(\phi^m - s\ell^m)$, respectively.

4. *(iv)* No agreement is CPNE only when there is a large degree of diseconomies of scale and market size asymmetry is sufficiently large.

4.2 Open Multilateralism

We now explore equilibrium analysis for open multilateralism, where trade agreements are open for all countries to join and provide equal opportunities for all countries, regardless of their market size, to encourage global free trade. Under an open multilateral approach to trade liberalization, Country s makes an announcement, denoted by α_s^o. Its strategy set, α_s^o, is defined as follows:

$$\alpha_s^o = \{\{\phi^o\}, \{M_m^o\}\}$$

where $\alpha_s^o = \{\phi^o\}$ indicates that country s does not announce to form a trade agreement with any country. Conversely, $\alpha_s^o = \{M_m^o\}$ denotes that country s announces to participate in a non-discriminatory, multilateral trade agreement.

It is important to note that the results observed within an open multilateral framework can correspond with those found in gated multilateralism and the patterns inherent in Nash equilibria. A key distinction in the stable Nash equilibrium under open multilateralism is that $\{s\ell^{m^o}\}$ never emerges as the CPNE among the three potential bilateral announcement profiles: $\{sm^{m^o}\}$, $\{s\ell^{m^o}\}$, and $\{m\ell^{m^o}\}$. This exclusivity is attributed to country ℓ's increased inclination for unilateral deviation in its announcement profile, opting for no agreement, driven by its desire to leverage its substantial market power without committing to reciprocal concessions to the smallest country s.

The nature of open multilateralism allows a significant market player to exert influence and capitalize on collective benefits while evading the mutual responsibilities typically mandated by more structured agreements. Aside from this, all CPNE conditions mirror those in gated multilateralism. Figure III illustrates the CPNE regions in the context of open multilateralism, shedding light on these stable Nash equilibria.

Now, we conclude the stability of the equilibrium trade regimes for our game of open multilateralism in the following proposition.

Proposition 5: Given the degree of economies of scale δ, the following holds:

(i) Global free trade is the unique CPNE unless market size asymmetry is large (global free trade is more likely as the degree of economies of scale rises) when $\alpha_\ell < \alpha_\ell(sm^{m^o} - F^{m^o})$.

(ii) When (F^{m^o}) is out of reach, (sm^{m^o}) and $(m\ell^{m^o})$ are CPNE only when the degree of market size asymmetry is at a medium range, i.e., $\alpha_m(\phi^{m^o} - sm^{m^o}) > \alpha_m > a_l(\phi^{m^o} - s\ell^{m^o})$ and $a_l(\phi^{m^o} - m\ell^{m^o}) > \alpha_\ell > a_l(\phi^{m^o} - s\ell^{m^o})$, respectively.

(iii) No agreement is CPNE only when there is a large degree of diseconomies of scale and market size asymmetry is sufficiently large.

5 Free Trade: Bilateralsim vs Multilateralism:

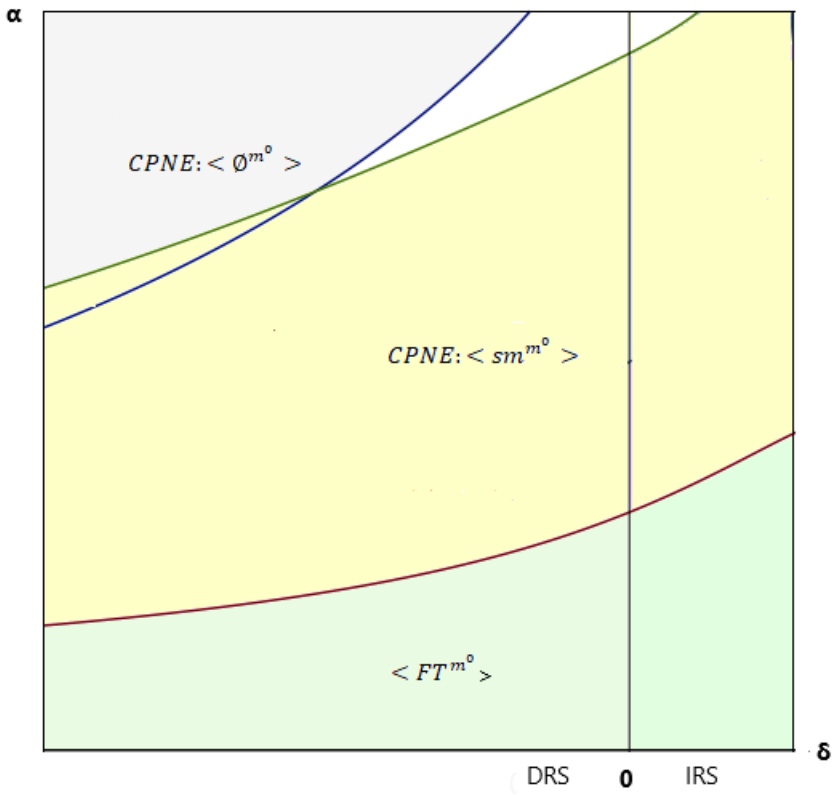

Figure-II: CPNE Multilateralism(Open)

In this section, we compare Propositions 2, 4, and 5 to address the following questions: *(i)* Which game presents a larger free-riding incentive? *(ii)* Is Article XXIV essential for achieving global free trade? *(iii)* If the answer to question *(ii)* is affirmative, then under what degree of economies of scale does this outcome persist? These inquiries help us comprehend the role of Article XXIV in the prospects for global free trade. We aim to confront Bhagwati's (1991) pivotal query: "Are PTAs building or stumbling blocks?" by adding a dimension considering economies of scale. This approach emphasizes the consequences of economies of scale in the pursuit of global free trade, offering a unique perspective and contribution to the existing body of literature.

Part *(i)* of each proposition indicates that Global Free Trade is the unique CPNE unless market size asymmetry is significantly large. Figure-III delineates the regions where Global Free Trade is attainable under two distinct games: Bilateralism and Multilateralism, encompassing both gated and open scenarios. It's crucial to observe that $\alpha_\ell(sm^m - F^m) = \alpha_\ell(sm^{m^o} - F^{m^o})$. The area beneath $\alpha_\ell(sm^m - F^m)$ represents a common zone where global free trade is feasible under both frameworks. Conversely, the space situated above this and below the IC: $\alpha_\ell(sm -$

F) highlights the circumstances under which global free trade is achievable solely through Bilateralism. Significantly, the CPNE conditions for Global Free Trade in both strategic games are defined by country ℓ's unilateral incentive to deviate from free trade agreements, influenced by its tendency toward free-riding. This behavior is observable in both contexts, yet Figure-III indicates that the point at which country ℓ engages in free-riding surfaces more promptly in the Multilateralism game:

$$\alpha_\ell(sm^m - F^m) = \alpha_\ell(sm^{m^o} - F^{m^o}) < \alpha_\ell(sm - F), \forall \delta. \tag{23}$$

The incentive for free riding is more pronounced under multilateralism, primarily because the largest country does not face any discrimination. Interestingly, facing discrimination can, in some contexts, actually expedite the achievement of global free trade. It creates a situation where there are consequences for not participating in mutual agreements, which, paradoxically, can encourage more inclusive trade practices and policies among countries. This dynamic underscores the necessary condition of Article XXIV in the broader pursuit of global free trade. Article XXIV indirectly addresses the issue of free riding. In this light, we advocate that Article XXIV plays an essential role in achieving global free trade, acting as a strong building block.

To analyze the role of economies of scale in the achievement of Global Free Trade, we refer to Figure-III. This achievement is likely when economies of scale are present. The left side of the figure shows that the area where Global Free Trade is attainable is very large with economies of scale. However, as we move to the right side of Figure-III, indicating a degree of diseconomies of scale, this area decreases. With economies of scale, the market power of large countries becomes less significant due to the diminishing returns on additional production, unless the parameter α is very high. This scenario implies that country ℓ is very large, and country s is very small. In this case, the return for country ℓ is reduced by providing access to smaller markets within its own market, thereby triggering the exercise of free-riding incentives. Conversely, with diseconomies of scale, as the degree of diseconomies increases, meaning the cost of production per unit rises, Global Free Trade becomes less attractive for the large country (ℓ). The reason is that country ℓ has its own large market, and the increasing cost of production may not be profitable for its producers in the largest market under Global Free Trade. Consequently, the incentive for free riding is exercised earlier under diseconomies of scale compared to when it is exercised under economies of scale.

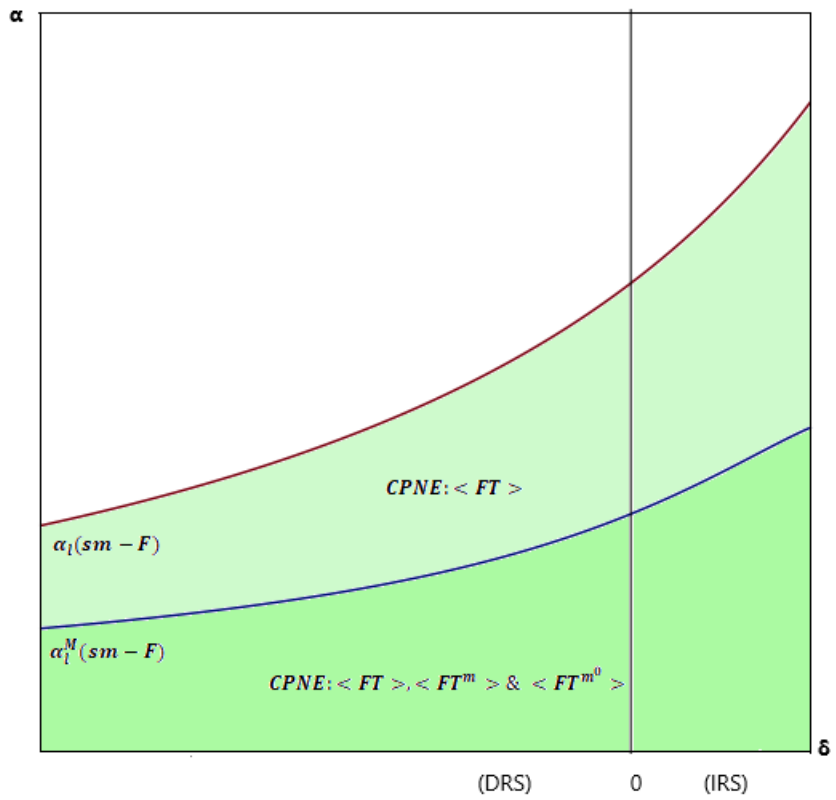

Figure-III: FT- Bilateralism vs Multilateralism

6 Conclusions:

In most of the literature on Preferential Trade Agreements (PTAs), the assumption of Constant Returns to Scale (CRS) production technology is prevalent. This assumption imposes a significant limitation on the analysis and may explain the ambiguity in existing studies regarding the effects of PTAs on trade, tariffs, and welfare. To address this limitation, we focus on the implications of economies of scale. On the other hand, Preferential Trade Agreements (PTAs) have become increasingly popular. In this context, we specifically examine Free Trade Agreements (FTAs) and investigate whether global free trade would be easier to achieve if countries pursued trade liberalization solely on a multilateral basis. We also analyze how eliminating the bilateral approach affects equilibrium outcomes. This exploration ties into Bhagwati's (1991) seminal question: *Are PTAs building or stumbling blocks?* Our focus is on highlighting the role of Article XXIV (Bilateralism) in the pursuit of global free trade.

In this paper, we evaluate the added value of Article XXIV in the presence of economies of scale and market size asymmetry. Ultimately, we argue that the economy of scale biases toward FTA (Bilateralism) formation, and that Article XXIV can act as a building block, playing a

crucial role in achieving global free trade. We consider a game of endogenous trade agreements involving three countries, where each country can choose to pursue either (a) bilateral, (b) multilateral, or (c) no trade liberalization. By applying this approach in the context of economies of scale and market size asymmetry, we analyze the discussion of building versus stumbling blocks and advance the discourse. Specifically, we advocate for Bilateralism over multilateralism (gated) for achieving global free trade, unless market size asymmetry is sufficiently large. FTAs act as strong building blocks when market size asymmetry is not sufficiently large. We conclude that Bilateralism may lead to global free trade in cases where multilateralism may not. When global free trade fails to materialize, the next agreement typically involves small and medium countries as a CPNE. This outcome is due to the unilateral deviation of the large country ℓ to exploit free-riding incentives. We discuss all possible cases where global free trade fails to emerge under both games. Our analysis also sheds light on tariff complementarity: the formation of a bilateral FTA induces each member to lower its tariff on the non-member country relative to the no trade agreement case. Our findings suggest that the tariff complementarity effect becomes more pronounced as the degree of economies of scale increases. Higher degrees of tariff complementarity result in less discrimination, and vice versa. We find that tariff complementarity holds for very high or low degrees of economies of scale. This contributes to addressing mixed empirical results on tariff complementarity in existing literature. Since the degree of tariff complementarity is industry-specific, it can be determined for both economies of scale and diseconomies of scale. Our future work aims to empirically test whether tariff complementarity holds across all degrees of economies of scale.

7 Appendix : Proofs

We impose restrictions on the parameters, $\underline{\delta} < \delta < \overline{\delta}$ and $0 \leq \alpha < \overline{\alpha}$, to guarantee positive export volumes under non-prohibitive tariffs. Using countries' welfare and optimal tariffs under the given trade regimes, we shall prove the following lemmas and propositions.

Proof of Lemma 1:

Lemma 1 directly follows from the welfare comparison of all countries for all α, for all δ.

Proof of Lemma 2:

(i) *Case 1 (Bilateral FTA $\langle sm \rangle$):* Country m will unilaterally deviate from the announcement profile $\sigma_m = \langle s, \phi \rangle$ to $\sigma_m = \langle \phi, \phi \rangle$ if and only if $\alpha_m > \alpha_m(\phi - sm)$, since in this case:

$W_m(\phi) > W_m(sm)$. Similarly, we can obtain announcement profiles supporting unilateral deviation from bilateral FTAs $\langle sl \rangle$ and $\langle ml \rangle$ to no agreement for country l, provided that α_l is sufficiently high in each case.

Case 2 (Hub-and-spoke agreement): Suppose country ℓ is the hub and countries s, m are spokes. Country ℓ will unilaterally deviate from the announcement profile $\sigma_\ell = \langle s, m \rangle$ to $\sigma_\ell = \langle \phi, \phi \rangle$ if and only if $\alpha_\ell > \alpha_\ell(\phi - \ell h)$, since in this case: $W_\ell(\phi) > W_\ell(\ell h)$. Similarly, we can obtain announcement profiles supporting the unilateral deviation of the larger country from other hub-and-spoke configurations for sufficiently high α.

Case 3 (Free Trade Agreement): Suppose all three countries are in a free trade agreement F. Country ℓ will unilaterally deviate from the announcement profile $\sigma_\ell = \langle s, m \rangle$ to $\sigma_\ell = \langle \phi, \phi \rangle$ if and only if $\alpha_\ell > \alpha_\ell(sm - F)$, since in this case: $W_\ell(sm) > W_\ell(F)$.

(ii) We found that for sufficiently high market size α, across all degrees of economies of scale δ, exclusion incentives arise for countries m and ℓ to exclude country s. Specifically, countries m and ℓ will jointly deviate from announcement profiles $\delta_m = \langle s, \ell \rangle$ to $\delta_m = \langle \phi, \ell \rangle$ and $\delta_\ell = \langle s, m \rangle$ to $\delta_\ell = \langle \phi, m \rangle$ respectively, when $\alpha_m > \alpha_m(m\ell - F)$, since: $W_z(m\ell) > W_z(F)$, for $z = m, \ell$.. This implies that countries m and ℓ prefer the bilateral FTA denoted $(m\ell)$ over the Free Trade Agreement (F), and thus have incentives to exclude country s. However, Lemma 1(ii) demonstrates that for all market sizes α and economies of scale δ, country s always prefers participation in the trade agreement over being excluded from agreement $(m\ell)$.

Nevertheless, due to the flexibility of FTAs, countries ℓ or m can jointly deviate with country s from their bilateral agreements to become hub countries:
$W_m(mh) > W_m(m\ell)$ and $W_s(mh) > W_s(m\ell)$, when $\alpha_m < \alpha_m(m\ell - mh)$, or
$W_\ell(\ell h) > W_\ell(m\ell)$ and $W_s(\ell h) > W_s(m\ell)$, when $\alpha_\ell < \alpha_\ell(m\ell - \ell h)$.
Hence, due to the flexibility of FTA formation, while exclusion incentives for countries m and ℓ exist, these incentives remain unexercised. Essentially, each member country under an FTA has limited ability to manipulate its terms of trade compared to the free trade scenario, which precisely encapsulates Lemma 2(ii).

(iii) To analyze the free-riding effects, we draw upon Lemma 2(i), which illustrates the crucial role that country ℓ, due to its relative size and influence, plays in the formation of global free trade. This influence is particularly evident when considering the condition:

$W_\ell(sm) > W_\ell(F)$, when $\alpha_\ell > \alpha_\ell(sm-F)$. Furthermore, no country in a bilateral agreement prefers to become a spoke country. Thus, no further deviation is possible, and the initial unilateral deviation from Free Trade $\langle F \rangle$ to the bilateral FTA $\langle sm \rangle$ (as detailed in the announcement profile in Case 3 of the previous proof) is self-enforcing.□

Proof of Proposition 1: The result follows immediately by comparing the countries' welfare functions under optimal tariff levels, together with Lemmas 1 and 2.□

References

[1] Ackerberg, D., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411–2451.

[2] Aghion, P., Antràs, P., & Helpman, E. (2007). Negotiating Free Trade. *Journal of International Economics*, 73, 1–30.

[3] Antweiler, W., & Trefler, D. (2002). Increasing returns and all that: A view from trade. *American Economic Review*, 92(1), 93–119.

[4] Azeem, M., Stoyanov, A., & Yildiz, H. M. (2024a). Role of Economies of Scale for the External Tariffs and Trade Agreement Formation: A Theoretical and Empirical Investigation. Working Paper.

[5] Azeem, M., Stoyanov, A., & Yildiz, H. M. (2024b). Empirical Evidence on Economies of Scale and External Tariffs: A Companion Analysis. Working Paper.

[6] Bagwell, K., & Staiger, R. W. (1997a). Multilateral Tariff Cooperation during the Formation of Free Trade Areas. *International Economic Review*, 38(2), 291–319.

[7] Bagwell, K., & Staiger, R. W. (1997b). Tariff revenue games. *Journal of International Economics*, 42, 1–30.

[8] Bagwell, K., & Staiger, R. W. (1998). The simple economics of labor standards and the GATT. *Journal of International Economics*, 46, 291–310.

[9] Bagwell, K., & Staiger, R. W. (1999). An Economic Theory of GATT. *American Economic Review*, 89(1), 215–248.

[10] Bagwell, K., & Staiger, R. W. (2005a). Enforcement and private political pressure. *Journal of Legal Studies*, 34, 471–515.

[11] Bagwell, K., & Staiger, R. W. (2005b). Multilateral trade negotiations and bilateral opportunism. *Journal of International Economics*, 67, 268–294.

[12] Bartelme, D., Costinot, A., Donaldson, D., & Rodriguez-Clare, A. (2024). The Textbook Case for Industrial Policy: Theory Meets Data. *Journal of Political Economy*, forthcoming.

[13] Basu, S., & Fernald, J. (1997). Returns to scale in U.S. production. *Journal of Political Economy*, 105(2), 249–283.

[14] Bernheim, B. D., Peleg, B., & Whinston, M. (1987). Coalition-proof Nash equilibria. *Journal of Economic Theory*, 42, 1–12.

[15] Bhagwati, J. (1991). *The World Trading System at Risk*. Princeton University Press.

[16] Bohara, A., Gawande, K., & Sanguinetti, P. (2004). Trade diversion and declining tariffs: Evidence from Mercosur. *Journal of International Economics*, 64(1), 65–88.

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

[18] Bond, E. (2009). The economics of preferential trade agreements. *World Economy*, 32, 725–742.

[19] Bown, C. P., Erbahar, A., & Zanardi, M. (2021). Global value chains and the removal of trade protection. *European Economic Review*, 140, 103937.

[20] Burnside, C. (1996). Production function estimation with aggregate shocks. Working Paper.

[21] Calvo-Pardo, H., Freund, C., & Ornelas, E. (2009). The ASEAN Free Trade Agreement: Impact on trade flows and external trade barriers. World Bank Policy Research Paper No. 4166.

[22] Chi, E. C., & Yildiz, H. M. (2024). Dynamic Preferential Trade Agreement Formation and the Role of Political Economy. *Economic Inquiry*, 62(2), 748–781.

[23] Chi, E. C., & Yildiz, H. M. (2024). Bilateralism vs. Multilateralism: Role of Political Economy for the Prospect of Multilateral Free Trade. *Canadian Journal of Economics*, 57(3), 870–900.

[24] Estevadeordal, A., Freund, C., & Ornelas, E. (2008). Does regionalism affect tariffs toward non-members? *Quarterly Journal of Economics*, 123(4), 1531–1575.

[25] Furusawa, T., & Konishi, H. (2007). Free trade networks. *Journal of International Economics*, 72, 310–335.

[26] Gandhi, A., Navarro, S., & Rivers, D. (2013). Identification and inference in nonlinear models using control functions. *Econometrica*, 81, 1883–1925.

[27] Gao, W., & Kehrig, M. (2017). Returns to Scale, Productivity and Competition: Empirical Evidence from U.S. Establishments. *SSRN Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2731596.

[28] Goyal, S., & Joshi, S. (2006). Bilateralism and Free Trade. *International Economic Review*, 47, 749–778.

[29] Grieco, P. L. E., Li, S., & Zhang, H. (2016). Production Function Estimation with Unobserved Input Price Dispersion. *International Economic Review*, 57(2), 665–690.

[30] Grossman, G., & Helpman, E. (1994). Protection for Sale. *American Economic Review*, 84, 833–850.

[31] Grossman, G., & Helpman, E. (1995). Trade wars and trade talks. *Journal of Political Economy*, 103(4), 675–708.

[32] Hall, R. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy*, 96, 921–947.

[33] Hall, R. (1990). Invariance properties of Solow's productivity residual. *Growth Economics*, 22, 229–244.

[34] Hillman, A., & Ursprung, H. (1996). The political economy of trade liberalization. *European Economic Review*, 40, 783–794.

[35] Konishi, H., & Furusawa, T. (2007). Trade networks revisited. *Journal of Economic Theory*, 72, 310–335.

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

[37] Krugman, P. (1991). Is bilateralism bad? In E. Helpman & A. Razin (Eds.), *International Trade and Trade Policy*. MIT Press.

[38] Lake, J. (2017). Trade agreements and welfare. *AEJ: Microeconomics*, 9, 299–328.

[39] Lake, J., & Krishna, K. (2019). Preferential trade agreements. *Journal of International Economics*, 117, 103–122.

[40] Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70, 317–341.

[41] Limão, N. (2006). Preferential trade agreements as stumbling blocks for multilateral liberalization. *American Economic Review*, 96, 896–914.

[42] Maggi, G. (2014). International Trade Agreements. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics*, Vol. 4, pp. 317–390. Amsterdam: North-Holland.

[43] Missios, P., Saggi, K., & Yildiz, H. M. (2016). External trade diversion and the nature of preferential trade agreements. *Journal of International Economics*, 99, 105–119.

[44] Ornelas, E. (2005a). Trade Creating Free Trade Areas and the Undermining of Multilateralism. *European Economic Review*, 49(7), 1717–1735.

[45] Ornelas, E. (2005b). Endogenous Free Trade Agreements and the Multilateral Trading System. *Journal of International Economics*, 67, 471–497.

[46] Ornelas, E. (2007). Exchanging market access at the non-member's expense. *Canadian Journal of Economics*, 40(1), 207–224.

[47] Riezman, R. (1999). Can bilateral agreements help induce free trade? *Canadian Journal of Economics*, 32, 751–766.

[48] Saggi, K. (2006). Preferential trade agreements and multilateral tariff cooperation. *International Economic Review*, 47, 111–136.

[49] Saggi, K., Stoyanov, A., & Yildiz, H. M. (2013). Multilateral trade negotiations and preferential trade agreements. *Journal of International Economics*, 89, 110–122.

[50] Saggi, K., & Yildiz, H. M. (2006). Bilateral Trade Agreements and the Feasibility of Multilateral Free Trade. *Review of International Economics*, 14(2), 356–373.

[51] Saggi, K., & Yildiz, H. M. (2009). External tariff determination under preferential trade agreements. *Journal of International Economics*, 79, 143–159.

[52] Saggi, K., & Yildiz, H. M. (2010). Bilateralism, Multilateralism, and the Quest for Global Free Trade. *Journal of International Economics*, 81, 26–37.

[53] Saggi, K., & Yildiz, H. M. (2011). Bilateral Trade Agreements and Feasibility of Multilateral Free Trade. *Review of International Economics*, 19, 356–373.

[54] Stoyanov, A., & Yildiz, H. M. (2015). Preferential vs multilateral trade liberalization and political economy. *European Economic Review*, 80, 140–164.

[55] Wooldridge, J. (2009). On estimating production functions. *Economics Letters*, 104, 112–115.

[56] Yanase, A., & Tsubuku, T. (2022). External tariffs under preferential trade agreements. *Journal of International Economics*, 137, 103–133.

[57] Yi, S. (1996). Endogenous formation of customs unions under imperfect competition: Open regionalism is good. *Journal of International Economics*, 41, 153–177.