Role of Economies of Scale for the External Tariffs and Trade Agreement Formation: A Theoretical and Empirical Investigation

Muhammad Azeem* Andrey Stoyanov† Halis Murat Yildiz‡

January 18, 2026

Abstract

In this paper, we provide a comprehensive theoretical and empirical analysis of the role of economies of scale on external tariff setting following the formation of free trade agreements (FTAs). We build an oligopoly model of trade with flexible degrees of economies of scale and show that the member countries have an incentive to reduce their external most favoured nation (MFN) tariffs following the FTA formation regardless of the degree of economies of scale. We further show that such effect strengthens as the degree of economies of scale rises. We empirically test these predictions and confirm that post-FTA tariff adjustments are consistent with our theoretical expectations, highlighting the importance of incorporating the economies of scale in the FTA analysis. Next, we develop a three-country dynamic FTA formation model to examine how the degree of economies of scale shapes the welfare implications of FTA formation and countries' incentives to enter into such agreements. Our findings suggest that the equilibrium path of FTA formation is fundamentally shaped by the degree of economies of scale in production technology: when countries exhibit strong preference for short-term gains, the FTAs are more likely to lead to global free trade with increasing or decreasing economies of scale than with constant return to scale.

Keywords: Economies of Scale, Tariff Complementarity, Tariff Bindings, Free Trade Agreement

JEL Classifications: C72, F12, F13

*Department of Economics, Toronto Metropolitan University. Email: ma122@torontomu.ca
†Faculty of Liberal Arts and Professional Studies, York University. Email: andreyst@yorku.ca
‡Department of Economics, Toronto Metropolitan University. Email: hyildiz@torontomu.ca

1 Introduction

Following the recent rapid proliferation of Free Trade Agreements (FTAs) globally, understanding their economic implications has become essential for policymakers. The economic benefits of FTAs are largely determined by how industries respond to trade liberalization, particularly through the impact of FTA tariff cuts on trade flows within the agreement and with non-member countries. The impact on trade flows, in turn, is closely tied to the presence of economies of scale in the production technology. For example, having a preferential access in an industry with economies of scale would lead to a greater expansion of trade within an FTA at the expense of trade from non-member countries, leading to change in incentives for optimal trade policies. Despite the theoretical importance of economies of scale in determining the trade and trade policy effects of FTAs, no study has yet systematically examined how economies of scale shape the properties and outcomes of FTAs.

In this study, we provide a comprehensive theoretical and empirical analysis of the role of economies of scale within the framework of FTAs. First, using a theoretical model we examine the external trade policy of FTA members and this analysis generates testable predictions on how FTAs affect import tariffs in different industries, depending on the presence and degree of economies of scale. Next, we empirically test these predictions and confirm that tariff adjustments following the FTA formation are consistent with theoretical expectations, highlighting the importance of considering the economies of scale in the FTA analysis. Building on this insight, we then develop a model of a dynamic FTA formation to explicitly examine how economies of scale shape countries' incentives to enter FTAs, and the welfare implications that arise from such agreements. We show that the equilibrium path of FTA formation is fundamentally shaped by the degree of economies of scale in production technology since it changes both short-run and long-run incentives of countries to form FTAs.

Our analysis begins by examining how economies of scale influence trade policies within FTAs, as tariffs changes following agreement formation are crucial for understanding the agreements' impact on trade and welfare of member and non-member countries. Theoretical and empirical studies remain divided on whether FTAs promote or hinder trade liberalization with non-member countries; however, both sides of this debate typically overlook the role of economies of scale in shaping FTA-related trade policy decision. To provide a more comprehensive perspective, we develop a theoretical model of FTA trade policy that explicitly accounts for the economy of scale. It is well known that internal economies of scale are incompatible with perfectly competitive markets and are typically a defining feature of models with monopolistic competition, where economies of scale arise from fixed costs of production in the presence of constant marginal costs. To allow greater flexibility in the type and degree of economies of scale, we rely on an oligopoly model of trade in which economy of scale affects the marginal and average costs of production. This paper, to the best of our knowledge, is the first to incorporate variable degrees of economies of scale to analyze both the theoretical and empirical implications for external tariff setting and the equilibrium path of FTA formation in a dynamic FTA formation setting.

In contrast to models with constant returns to scale, where domestic and foreign markets are independent, the economies or diseconomies of scale in our model create strategic interdependence between a firm's domestic and export sales. Our first theoretical finding demonstrates the robustness of the tariff complementarity effect, wherein a reduction in the FTA internal tariffs among member countries leads a reduction in the welfare-maximizing external tariffs on non-members. This effect persists regardless of the sizes of FTAs, the levels of preferential tariffs, and the degree of economies of scale. This is primarily due to the fact that the preferential market access among FTA members and the discrimination faced by non-member firms reduce the significance of non-members as a source of rent extraction, thereby incentivizing members to set lower external tariffs.

Then, we turn our attention to examine how the degree of economies of scale affects the tariff complementarity effect across FTAs of varying sizes and levels of preferential market access. From the model, we derive a simple reduced-form relationship between changes in FTA external tariffs in a sector and its economies of scale. For a given tariff preference, an increase in FTA size and in the economies of scale within an industry amplify the tariff complementarity effect, further reducing member countries' incentives to impose external tariffs compared to industries with constant or diminishing returns to scale. This amplification occurs because economies of scale increase efficiency gains from expanded intra-FTA output and trade, reducing the relative share of non-members' exports to the FTA and diminishing members' incentives for protectionism. As a result, the model predicts a stronger tariff complementarity effect in industries with economies of scale compared to industries with diseconomies of scale.

We test this prediction by analyzing the response of MFN tariffs to preferential tariff reductions, using product-level tariff data across all FTAs signed from 1989 to 2014. To capture the dynamics and evaluate the persistence of the tariff complementarity effect, we estimate the evolution of MFN tariff adjustments over six years following FTA implementation. Working with the large sample of diverse FTAs, we have to address the potential endogeneity of the FTA preference margin, which may arise if economic or political shocks influence both FTA tariff reductions and MFN tariff adjustments. To address this, we employ an instrumental variable approach developed by Stoyanov (2024) [9], using tariff preferences from a country's first FTA in our sample as an instrument for tariff preferences in subsequent agreements, provided that the initial FTA was implemented at least three years prior. This strategy mitigates the risk of contemporaneous shocks contaminating our estimates. In order to categorize industries into increasing returns to scale and decreasing returns to scale, we use the estimates of the return to scale parameters from the two recent studies by Bartelme et al. (2024) [7] and Lashkaripour and Lugovskyy (2023) [27]. If an industry is identified as exhibiting increasing or decreasing returns to scale in both studies, we classify it accordingly in our analysis.

In line with the predictions of the model, we find that while the tariff complementarity effect is observed across all industries regardless of their economies of scale, with MFN tariffs adjusting steadily over the first three years of FTA implementation in response to preferential tariff reductions. More importantly, industries characterized by increasing returns to scale show markedly greater reductions in MFN tariffs, a difference that is both statistically and economically significant. For instance, while a one percentage point increase in the FTA preferential tariff margin yields less than a 0.01 percentage point decrease in MFN tariffs for industries in the lowest quartile of economies of scale, this response is more than four times greater for industries in the top quartile.[1] These results indicate the significance of economies of scale in FTA trade policy decisions, suggesting that policymakers in various countries take this aspect of industry's technology into account when designing optimal trade policies.

Having established the importance of economies of scale for FTA trade policies, we next extend our theoretical model to examine how this feature of technology shapes both the welfare implications of FTA formation and countries' incentives to participate. Specifically, building on Lake and Yildiz (2016) [24] and Lake et al. (2020) [26], we construct a dynamic model of FTA formation with three symmetric countries, incorporating flexible production technoilogy

[1] Our elasticity estimates are generally lower than those reported in previous studies, likely due to our inclusion of all FTAs, many of which are between countries with modest trade volumes. When applying our approach to large FTAs, such as NAFTA, we observe elasticity estimates between 0.1 and 0.25, which aligns more closely with findings from studies focused on major trade agreements.

into oligopolistic markets to examine how varying degrees of economies of scale influence the equilibrium path of FTA formation. In this dynamic setting, countries can form only one FTA per period, within a context where internal FTA tariffs are fully eliminated.

The primary focus of this model lies in examining how different levels of economies of scale affect the likelihood of expanding FTAs to global free trade. This analysis hinges on two key factors. First, whether countries already in an FTA have a strategic interest in excluding an outsider country (the *exclusion incentive*). Second, whether the non-member country has an incentive to join the FTA or to remain outside and free-ride on tariff reductions by the member countries (the *free-riding incentive*). The balance between these forces determines the potential for FTA to expand and evolve towards the global trade liberalization. Another essential component of our dynamic model is the distinction between short-run and long-run analysis. Specifically, incentives based on payoff comparisons between the current period and the immediate next period are called *myopic incentives*, as they primarily capture short-run considerations. In contrast, when countries compare payoffs in the current period to periods beyond the immediate next, these are referred to as *farsighted incentives*, which involve long-run considerations.

Our analysis is structured to examine free-riding and exclusion incentives both in the short run and long run. Suppose two countries form an FTA while a third country remains an outsider. Myopically, the FTA outsider must decide whether to establish a separate FTA with one of the existing members or to remain outside the bilateral FTA, with the latter choice indicating that the free-riding incentive dominates. Farsightedly, however, such an FTA may represent a step toward global free trade, with the long-run policy choice involving a payoff comparison between global free trade and remaining outside the bilateral FTA. For the FTA insiders, we first examine whether any of them would benefit from entering into another FTA with the outsider myopically. Farsightedly, the formation of a second FTA is likely to lead to global free trade, requiring member countries to weigh the decision to exclude the non-member or to pursue global free trade through a sequence of bilateral FTAs. In this dynamic setting, the discount factor serves as a mediator between myopic and farsighted incentives for both FTA insiders and the outsider, determining whether dynamic free riding and/or dynamic exclusion incentives arise.

Examining the role of economies of scale, we find that the welfare gain from a bilateral FTA increases with the degree of economies of scale. This effect arises from the combined benefits of reciprocal preferential market access and efficiency gains generated by economies of scale, which yield greater benefits for FTA members. Upon forming the first FTA, its members always have a myopic incentive to pursue another FTA with the outsider. For the outsider, however, the decision to form a new FTA with one of the existing members involves a tradeoff of two opposing forces. On one hand, the outsider gains from tariff reductions by FTA members that arise from tariff complementarity. That tariff reduction becomes deeper with stronger economies of scale, thereby reducing the outsider's incentive to form additional agreements. On the other hand, the economies of scale increase the value of tariff-free access to a member's market, eliminating the discrimination the outsider faces in that market. Importantly, the outsider's gains from forming a second FTA with one of the original members are limited, as it merely equalizes tariffs on imported goods in that market. As such, we find that the outsider consistently holds a *myopic free-riding incentive*, hindering any further trade liberalization, and that incentive becomes stronger as economies of scale within the industry increase.

Turning to the long-run analysis, we find that the FTA outsider always has a farsighted incentive to participate in an FTA expansion process that ultimately leads to global free trade. This incentive strengthens as economies of scale increase, revealing a tension between the outsider's myopic and farsighted incentives: while myopically inclined to free-ride and avoid joining the existing FTA, the outsider farsightedly favors eliminating discrimination in both FTA insider markets through global free trade. For FTA insiders, in contrast, there is no farsighted incentive to permanently exclude the outsider from FTA expansion. Instead, insiders stand to benefit from the elimination of tariff barriers imposed by the outsider, especially as economies of scale amplify efficiency gains through expanded trade. Larger output, coupled with reduced average production costs, enhances the welfare gains from progressing toward global free trade. Thus, for insiders, myopic and farsighted incentives are aligned, preventing the emergence of dynamic exclusion incentives.

The above discussion implies that the expansion of a bilateral FTA toward global free trade is ultimately dependent on the outsider's dynamic free-riding incentives. In analyzing these dynamic free-riding incentives, we find that, relative to constant returns to scale, both the outsider's myopic incentive to free-ride and farsighted incentive to achieve global free trade increase with the degree of economies of scale. However, as economies of scale become stronger, the rise in the farsighted incentive for free trade dominates the myopic incentive to remain outsider, thereby raising the critical discount factor below which dynamic free-riding incentives emerge. Conversely, under diseconomies of scale, both myopic and farsighted incentives get weaker, though the myopic incentive falls more substantially, again leading to a higher critical discount factor. Consequently, our results indicate that the threshold discount factor, above which the first FTA expands toward global free trade, is a convex function of the degree of economies of scale, reaching its minimum at constant returns to scale. This finding is significant, as much of the literature assumes constant returns when assessing whether FTA formation acts as a building or stumbling block for attaining global free trade. Our findings suggest that the equilibrium path of FTA formation is fundamentally shaped by the degree of economies of scale in production technology, making global free trade less attainable under either economies or diseconomies of scale when countries are sufficiently impatient.[2]

Note that there is a large literature on the preferential trade agreement (PTA) formation and its implications on the prospects of multilateral free trade in a variety of setting (see Maggi (2014) [29] and Lake and Krishna (2019) [25] for excellent surveys). 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.[3] 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) [1]; Furusawa and Konishi (2007) [18]; Saggi and Yildiz (2010) [43] and Stoyanov and Yildiz (2015) [49]. Specifically, focusing on the FTA formation incentives, Saggi (2006) [42], Saggi and Yildiz (2011) [44], Ornelas (2005a) [37] and Stoyanov and Yildiz (2015) [49] 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. Finally, similar to the present paper, Lake and Yildiz (2016) [24], Lake et al. (2020) [26] and Chi and Yildiz (2024b) [15] employ a dynamic PTA formation model in examining the equilibrium path of PTAs. Unlike the present paper, they utilize perfectly competitive models with DRS. Since

[2]This finding stands in line with the static model of Saggi and Yildiz (2010) [43] who shows that free riding incentive determines the stability condition of global free trade while exclusion incentive does not get exercised in an FTA formation game. However, unlike the present paper, note that Saggi and Yildiz utilizes a static competing exporters model with endowment and asymmetric endowment is needed for free riding incentive to arise.

[3]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) [21], Bhagwati (1991) [8], Yi (1996) [51], Saggi (2006) [42], Bagwell and Staiger (1997a, 1997b, 1998, 2005a, 2005b) [2, 3, 4, 5, 6], Krishna (1998) [20], Riezman (1999) [41], Bond et al. (2004) [10], Bond (2009) [11], Goyal and Joshi (2006) [19], Furusawa and Konishi (2007) [18], Aghion et al. (2007) [1], Missios et al. (2016) [32], Ornelas (2005a, 2007) [37, 38], Saggi et al. [45, 46, 47], Stoyanov and Yildiz (2015) [49], Lake (2017) [23], Lake et al. (2020) [26], Chi and Yildiz (2024a, 2024b) [14, 15].

the welfare properties of FTAs, and thus the incentives of countries to form FTAs heavily rely on the scale economies of production technology, this paper aims at filling the gap in the literature via incorporating flexible degree of economies of scale. The objective is to gain a better understanding of the equilibrium path of FTAs and whether FTA formation undermines multilateral trade liberalization in a more general setting.

This paper contributes to the existing literature on trade agreements in three main ways. First, we show that the tariff complementarity effect is robust regardless of the size of an FTA, levels of preferential margin, and the degree of economies of scale. The term of tariff complementarity was first coined by Bagwell and Staiger (1997a) [2] and has since been supported by various theoretical studies, including those by Saggi and Yildiz (2010) [43], Saggi et al. (2013) [45] and Ornelas (2005a, 2007) [37, 38]. Unlike the present study, these theoretical papers assume away the role of economies of scale in the external trade policy setting. From an empirical perspective, Estevadeordal et al. (2008) [16], Calvo-Pardo et al. (2009) [13] and Mai and Stoyanov (2015) [30] give support for the existence of tariff complementarity.[4] We confirm these findings and go one step further and examine how the nature and the degree of economies of scale affect the external tariffs of the FTA member countries. Our results imply that scale economies play pivotal roles in FTA trade policy decisions, suggesting that policymakers should take it into account when designing optimal trade policies. Second, while there is a large literature on the static and dynamic preferential trade agreement (PTA) formation that examines the implications of PTAs for the prospects of multilateral free trade in a variety of settings, the roles of the presence and degrees of economies of scale have been completely ignored.

Finally, similar to the present paper, Lake and Yildiz (2016) [24], Lake et al. (2020) [26] and Chi and Yildiz (2024b) [15] employ a dynamic PTA formation model in examining the equilibrium path of PTAs. Unlike the present paper, they utilize perfectly competitive models with DRS. Since the welfare properties of FTAs, and thus the incentives of countries to form FTAs heavily rely on the scale economies of production technology, this paper aims at filling the gap in the literature via incorporating flexible degree of economies of scale. The ultimate objective is to gain a better understanding of the equilibrium path of FTAs and whether FTA formation undermines multilateral trade liberalization in a more general setting.

[4] However, not all findings support this concept Limão (2006) [28] disputes it and provides a strong evidence for tariff substitutability in developed countries by showing that the EU and US increased their tariffs after forming PTAs.

2 Theoretical Model

Our theoretical framework builds upon an adapted version of the oligopoly trade model introduced by Brander and Krugman (1983) [12]. We consider n symmetric countries, indexed by i. There are two industries: a perfectly competitive numeraire industry, labeled y, and oligopolistic non-numeraire industry, labeled x. Firms in industry x engage in Cournot competition and produce homogeneous product.

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

$$U_i\left(x_i, y_i\right) = u_i\left(x_i\right) + y_i \tag{1}$$

where the subutility $u_i(x_i)$ is quadratic, defined as $u_i(x_i) = \alpha x_i - \frac{x_i^2}{2}$. The term x_i represents the total consumption of x in country i, which is the sum of consumption of domestically produced good and imports of x from all other countries: $x_i = \sum\limits_{z=1}^{n} x_{iz}$. This formulation leads to a linear demand function for product x in country i :

$$p_i = \alpha - x_i \tag{2}$$

where p_i denotes the price of good x in country i.

On the production side, the numeraire good y is produced by perfectly competitive firms operating under constant returns to scale technology, and is freely traded internationally. In contrast, good x is produced in an oligopolistic market, with one profit-maximizing firm operating in each country. Since there is only one firm in each country producing good x, we use the same index i to denote both the countries and their respective firms in industry x. All oligopolistic firms are symmetric and share the following quadratic cost function:

$$C_i\left(q_i\right) = cq_i - \delta q_i^2 \tag{3}$$

where q_i denotes the total production of good x by firm i, which is the sum of its sales across all countries $q_i = \sum\limits_{z=1}^{n} q_{iz}$.

The parameter δ in the cost function captures the extent of economies of scale:

$$\delta = \begin{cases} > 0 \text{ if there are Economies of Scale,} \\ = 0 \text{ if there are Constant Returns to Scale,} \\ < 0 \text{ if there are Diseconomies of Scale} \end{cases} \tag{4}$$

Most existing studies that use the oligopoly model of trade assume constant marginal cost, implying $\delta = 0$, which corresponds to the constant returns to scale technology. In contrast, we adopt a more flexible framework that allows for both increasing and decreasing returns to scale. We use the quadratic cost function for analytical tractability, though it is important to note that this formulation produces results equivalent to those obtained from a general Cobb-Douglas production function. In fact, there is a one-to-one correspondence between the parameter δ in our cost function and the sum of the exponent parameters in the Cobb-Douglas production function. This relationship between the parameter δ in the cost function (3) and Cobb-Douglas parameters is discussed in more detail in the appendix.

2.1 Optimal outputs and tariffs

Let t_{iz} denote the specific tariff imposed by country i on imports of product x from country z, and let t_i represent the vector of tariffs that country i imposes on all its trading partners. Given the tariffs faced by firm i in all markets, the firm solves the following profit-maximization problem:[5]

$$\max_{x_{iz}} \pi_i(\mathbf{t}) = \sum_{z=1}^{n} p_z(\mathbf{t})x_{iz}(\mathbf{t}) - cq_i(\mathbf{t}) + \delta q_i^2(\mathbf{t}) - \sum_{z \neq i} t_{iz}x_{iz}(\mathbf{t}),\ i = 1, 2...., n \tag{5}$$

where $\mathbf{t}$ represents the tariff vector across all countries, involving external tariffs for each market. The Cournot Nash equilibrium for output and prices can be derived under any given trade policy regime, as shown in the appendix. It is important to note here that the presence of economies of scale creates a link between a firms' domestic and export market outputs. This interdependence of firms' strategies implies that, unlike under constant returns to scale, the markets are not strategically separated, and change in sales in one market affects the marginal cost of production for sales in all other markets.

Given the equilibrium prices and export volumes across all countries, the national welfare

[5]It is important to note that $t_{ii} = 0$ since countries do not impose tariffs on domestic firms, and there are no tariffs on the numeraire good.

of country i, denoted by $\mathbf{W}_i(\mathbf{t})$, is defined as the sum of consumer surplus (CS_i), producer surplus (PS_i), and tariff revenue (TR_i):[6]

$$\mathbf{W}_i(t_i) = \underbrace{u_i(x_i(t_i)) - p_i(t_i)x_i(t_i)}_{CS_i(t_i)} + \underbrace{\pi_i\,(t_i)}_{PS_i(t_i)} + \underbrace{\sum_{z \neq i} t_{iz}x_{zi}}_{TR_i(t_i)} \tag{6}$$

Using this welfare function, we can obtain welfare-maximizing tariffs under various trade regime. We are primarily interested in two trade regimes: one in which no trade agreements exist, and the other where a group of countries form a trade agreement.

In the absence of a trade agreement, each country independently sets its optimal non-discriminatory MFN tariff on product x, as required by Article I of the GATT. We refer to the trade regime without trade agreements as ϕ. Non-discrimination implies $t_{iz} = t_i(\phi)$ for all $z \neq i$. The optimal MFN tariff of country i is calculated as

$$t_i(\phi) \equiv \arg\max_i \mathbf{W}_i\,[\phi]$$

Next, consider a scenario where m countries form an FTA. Since all countries in the model are symmetric, adjusting m allows us to vary the size of the FTA relative to non-members. Naturally, the tariff complementarity effect is stronger in FTAs formed between major trade partners and weaker in FTAs involving countries with less trade, as we demonstrate below. The parameter m shows how the size of preferential trade between FTA partners affects the tariff complementarity.

Let M denote the set of member countries and K the set of non-member countries under this FTA. We assume that countries i and j are generic member countries ($i, j \in M$), and countries k and r are generic non-member countries ($k, r \in K$). We denote a trade regime under such FTA as f, under which member countries grant preferential access to each other by imposing a preferential tariff τ_i on intra-FTA trade. At the same time, each member country retains the ability to impose independent external MFN tariff $t_i(f)$ on imports from non-member countries, where $\tau_i < t_i(f)$. The optimal MFN tariff of a member country under trade regime f is obtained as:

$$t_i(f) \equiv \arg\max_i \mathbf{W}_i\,[f] \text{ subject to } t_{iz} = \tau_i \text{ for } z \in M \text{ and } z \neq i,$$

[6] We report all supporting calculations and proofs in the appendix.

while the non-member countries set their MFN tariffs according to:

$$t_k(f) \equiv \arg\max_{k} \mathbf{W}_k\,[f] \;\text{ for } k \in K$$

We now turn to the analysis of how three key factors – the size of an FTA (m), preferential tariffs (τ) and the degree of economies of scale (δ) – influence the external tariffs imposed by member countries on non-members following the transition from no agreement (trade regime ϕ) to the FTA (trade regime f). However, to understand the changes in external tariffs, we must first consider how FTA formation affects trade flows among members and non-members.

When an FTA is formed, production in member countries will increase in response to preferential access to each other's markets. The increase in production will be stronger in IRS industries and weaker in DRS industries. For the given preferential tariffs, this response is further amplified by the size of the FTA, as greater market access enables firms to lower their average costs more significantly in IRS industries, raising their mark-ups. In contrast, non-member countries face discrimination in members' markets and do not benefit from the economies of scale as much as the FTA members enjoy. As a result, the market shares of FTA member firms relative to non-member exporters rises in FTA member countries, particularly in industries with stronger economies of scale. To quantify the impact of FTA formation on external tariffs, let $\Delta t_i(\phi - f) = t_i(\phi) - t_i(f)$ denote the change in the MFN tariffs imposed by member country i on imports from non-member countries following the formation of an FTA. The resulting shift in production structure between member and non-member countries, driven by preferential market access, economies of scale, and the size of the FTA, leads us to the first main result of this section, summarized in the following proposition.

Proposition 1. *Suppose there are n countries and m of them form an FTA. Then the following holds:*

(i) Regardless of the size of the FTA (m) or the degree of economies of scale (δ), the tariff complementarity effect always holds for the given level of preferential tariffs (τ): $\Delta t_i(\phi - f) > 0$.

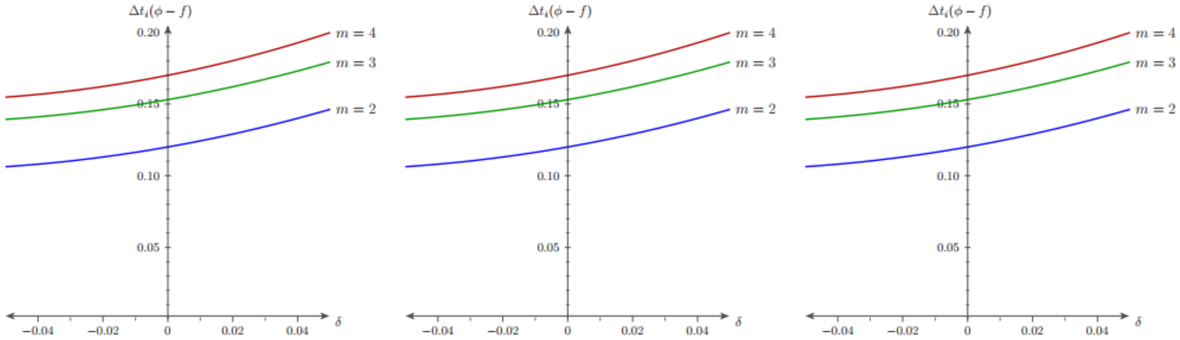
(ii) For the given level of preferential tariffs, the strength of the tariff complementarity effect increases with the the degree of economies of scale (δ).

(iii) For the given level of preferential tariffs (τ), the positive effect of the degree of economies of scale on the tariff complementarity gets stronger as the size of an FTA (m) rises,.

The first part of the proposition establishes that, irrespective of the size of the FTA, the level of preferential tariffs, or the degree of economies of scale, member countries have an incentive to reduce their external tariffs on imports from non-member countries following the formation of the FTA. The rationale behind this result lies in the preferential access between member countries, which diminishes the importance of non-member firms as a rent extraction source in the markets of FTA members, thus reducing the incentive to maintain high external MFN tariffs. As a result, the reduction in external tariffs on the imports from non-member countries $\Delta t_i(\phi - f) > 0$ consistently holds. This mechanism is known in the literature as the *tariff complementarity effect,* a term introduced by Bagwell and Staiger (1997) [**?**].

The second part of the proposition extends the analysis by examining how the rate of reduction in the external MFN tariffs of FTA members changes with the level of economies of scale. It states that, for a given size of the FTA and level of preferential tariffs, the presence of increasing returns to scale amplifies the tariff complementarity effect. The logic is as follows: economies of scale boost efficiency gains from increased intra-FTA output and trade, which reduces the relative share of non-members' exports to the FTA and further diminishes members' protection incentives. Consequently, a higher rate of reduction in the external MFN tariffs of FTA members occurs in industries with economies of scale compared to industries with CRS and those with diseconomies of scale.

The third part of the proposition examines the interaction between the FTA size and economies of scale on the change in the tariff complementarity effect for the given level of preferential tariffs. We find that an increase in FTA size strengthens the impact of economies of scale on the rate of external MFN tariff reduction of FTA members. Intuitively, as the FTA expands and FTA members impose lower internal tariffs on each other, intra-FTA production and trade increase further, boosting efficiency gains due to a higher degree of economies of scale. This would further reduce the relative share of non-members' exports to FTA members, making rent extraction for FTA members from FTA non-members less significant. As a result, FTA members have a greater incentive to lower their external tariffs on export from FTA non-members.



(a) $\tau = 0.01$ **(b)** $\tau = 0.03$ **(c)** $\tau = 0.05$

Figure 1: Tariff Complementarity Effect ($\alpha = 1, n = 5$)

Figure 1 illustrates the relationship between the change in external MFN tariffs of FTA members on exports from the FTA non-member and the degree of economies of scale for the given level of preferential tariffs and FTA size, providing a visual representation of the entire proposition outlined above. First, note that, for the given level of preferential tariffs, all curves, representing the reduction in external MFN tariffs of FTA members on non-members, remain above the horizontal axis, and thus the tariff complementarity exists regardless of the degrees of economies of scale and FTA size (part (i) of Proposition 1). Next, for any given size of the FTA and level of preferential tariffs, the positive slope of the curves indicates that the tariff complementarity effect becomes stronger as the degree of economies of scale increases (part (ii) of Proposition 1). Finally, for all levels of preferential tariffs, as the size of the FTA increases, the external MFN tariff reduction curves shift upward, reflecting the interaction effect described in part (iii) of Proposition 1.

The results presented above confirm the presence of the tariff complementarity effect in FTAs, which has been empirically validated in studies such as Estevadeordal et al. (2008) [16], Mai and Stoyanov (2015) [30], Kuenzel and Sharma (2021) [22], Limão (2006) [28] provides contradictory evidence in his analysis of U.S. tariffs. Our model confirms the existence of tariff complementarity within a more general framework that incorporates flexible economies of scale. Furthermore, the model offers new insights into how the degree of economies of scale influences the strength of the tariff complementarity effect in FTAs. These insights yield a testable hypothesis, which will be explored empirically in the following section.

3 Empirical Evidence

In this section, we present empirical evidence supporting the main mechanism in our theoretical model. Our analysis shows that industries with stronger economies of scale experience a more pronounced tariff complementarity effect, as preferential tariff reductions in FTAs result in more significant declines in the MFN tariffs. These findings confirm that economies of scale significantly influence FTA trade policies, motivating a deeper theoretical investigation in the next section of how scale economies affect the formation and structure of FTAs.

3.1 Empirical framework for estimating tariff complementarity

Estimating the impact of economies of scale on tariff complementarity requires a robust methodology for capturing the dynamics of tariff reductions following FTA formation. Most existing studies in the literature focus on individual FTAs, which limits the generalizability of their findings. By concentrating on a single FTA or a specific country, these studies often rely on context-specific factors that may not apply universally, thus restricting their relevance to broader trade policy discussions.

In this paper, we aim to analyze tariff complementarity across a wide range of FTAs using data from multiple agreements, thereby enhancing the reliability of our results and ensuring that our conclusions are not limited by the specifics of a single FTA. However, working with a large sample of FTAs introduces several methodological challenges as we discuss below. To navigate these issues, we build upon the framework developed by Bi and Stoyanov (2024) [9], whose recent study introduces a flexible approach for estimating tariff complementarity that can be applied broadly across different FTAs. In this section, we outline their methodology and explain how we adapt it to estimate the impact of economies of scale on the strength of tariff complementarity within FTAs.

Building on the work of Bi and Stoyanov (2024) [9], the empirical strategy for estimating the tariff complementarity effect is based on the following model:

$$\Delta_T MFN_{cni} = \beta_T \Delta_T PM_{cni} \times PS_{cni} + \mu_c + \mu_i + \mu_n + \varepsilon_{cin}. \tag{7}$$

In equation (7), the subscripts c and i denote country and product, respectively, while n denotes the FTAs index for country c. If a country has multiple FTAs in place, this setup treats each FTA of a country as a separate cluster of observations.

The dependent variable, $\Delta_T MFN_{cni}$, represents the change in the MFN tariff rate for product i by country c between the last year prior to the FTA n ($t = -1$) and year T after the FTA's implementation ($t = T$): $\Delta_T MFN_{cni} = MFN_{cni}^{t=-1} - MFN_{cni}^{t=T}$. The main explanatory variable is the product of two components. The first is $\Delta_T PM_{cni}$, which is the tariff preference margin in industry i, granted by country c under FTA n in year T. More formally, the tariff preference margin is constructed as $\left(MFN_{cni}^{t=-1} - PT_{cni}^{t=T}\right)$, which is the difference between country c's MFN tariff the last year before the FTA's commencement and its preferential tariff in year T of the FTA.[7] In this way, the variables $\Delta_T PM$, for $1 \leq T \leq 6$, capture the evolution of preferential tariff of FTA n over the first six years of the agreement.

The second term on the right-hand side in equation (7), PS_{cni}, represents the share of imports originating from countries eligible for tariff preferences under FTA n. Proposition 1 suggests that the tariff complementarity effect within an FTA is influenced not only by the level of tariff preferences, but also by the size of the agreement. Specifically, it implies that, for a given level of economies of scale in an industry, the strength of tariff complementarity increases as the size of the FTA grows and the preferential tariff falls. To align our empirical framework with the theoretical model, equation (7) interacts the tariff preference on the right-hand side with the share of imports that come from FTA partner countries. This interaction term allows us to capture the theoretical prediction that tariff complementarity will be stronger in FTAs between major trade partners, where the share of trade is higher, and weaker in FTAs involving countries with less trade. To minimize potential reverse causality problem, we use trade data from the last year before the FTA to construct PS_{cni} variable.

To summarize the discussion above, we expect that more extensive FTA tariff preferences granted within an FTA will result in deeper reductions in the MFN tariff, and more so in industries with larger presence of firms from FTA partner countries. Therefore, tariff complementarity between preferential and MFN tariffs imply $\beta_T > 0$. To capture the dynamic effects of preferential trade liberalization on MFN tariffs, we estimate model (7) over different time horizons ($1 \leq T \leq 6$), allowing us to identify how external tariff adjustments evolve in response to tariff preferences over the first six years of the agreement.

A key challenge in estimating the causal effect of FTA tariff preferences on MFN tariffs using model (7) is the potential endogeneity of the main explanatory variable. The preferential

[7] To mitigate concerns of simultaneity with the dependent variable, we use the MFN tariff from the last year before the FTA to construct the preference margin for subsequent years.

market share, PS_{cni}, may be endogenous due to reverse causality – where the initial MFN tariff level affects trade with FTA partner countries differently than with other countries or due to omitted variables that simultaneously influence both a country's import composition and its trade policy decisions. To address this, we employ an instrumental variable for PS_{cni}. This instrument is constructed as the total exports of FTA partner countries relative to the global trade value of the corresponding product (excluding exports to country c). This instrument is plausibly exogenous to country c's MFN tariff, as it is independent of its domestic market conditions, while also being relevant since it captures the comparative advantage of c's trade partners.

The second source of endogeneity stems from the potential endogeneity of the FTA preference margin variable, $\Delta_T PM_{cni}$. Although preferential tariffs are generally negotiated prior to FTA implementation and are thus pre-determined relative to subsequent MFN tariff changes, the time window between these negotiations and subsequent tariff adjustments is often too short to completely rule out co-determinants. Persistent economic or political shocks in country c may simultaneously affect the government's decision regarding both the pace of FTA tariff reductions and the adjustment of MFN tariffs shortly after the FTA takes place.

We address this challenge with the instrumental variable strategy as well. Specifically, we instrument $\Delta_T PM_{cni}$ with the tariff preference granted by country c on imports of product i under its first FTA in our database, provided that this initial FTA was formed at least three years prior to FTA n. Since the tariff preferences from the first FTA were implemented at least three years before subsequent agreements, this instrument remains unaffected by any socioeconomic shocks that may have affected later FTA negotiations. Furthermore, the instrument is highly relevant in the first stage, as it reflects country c's historical protectionist tendencies, which tend to favor certain industries regardless of current market conditions. On the downside, this strategy reduces the sample size by excluding each country's first FTA, along with any subsequent FTAs signed within three years of that first one, resulting in a loss of data for some countries and agreements.

3.2 Data

Estimating the heterogeneity of tariff complementarity across industries requires three sets of data: trade flows, tariff schedules, and measures of industry-specific economies of scale. Below, we describe how we construct the measure of economies of scale and describe the publicly available sources for trade and tariff data.

The data on trade flows, import tariffs, and FTA membership are obtained from Saggi et al. (2018) [46], who sourced the original data from the World Integrated Trade Solutions (WITS). The dataset spans the period from 1989 to 2014 and provides an unbalanced panel covering 192 importing countries and 253 exporting countries, and $5,413$ product categories, classified at the 6-digit HS level.

In our study, the parameter measuring industry's economy of scale plays a central role in the analytical framework. To obtain a measure of the economies of scale, we draw on two recent studies that provide estimates the return to scale by industry. The first study by Bartelme et al. (2024) [7] estimates the economy of scale parameter for 34 industries classified at 2-digis SIC level, using data on sector size and observed trade shares as a source of identifying variation. The second study by Lashkaripour and Lugovskyy (2023) [27] employs firm-level trade data from Colombia to estimate the return to scale for 14 sectors based on the ISIC Rev.3 classification. We map both estimates into 2-digit HS industry classification using WITS concordance tables.

It is important to note that there is a significant disparity between the return to scale estimates provided by Bartelme et al. (2024) [7] and Lashkaripour and Lugovskyy (2023) [27], with the negative sample correlation between the two sets of estimates. The parameters produced by both studies are likely subject to significant measurement error, and there is no objective test available to evaluate reliability of the return to scale parameter estimates.

To minimize the effect of measurement error on our analysis, we avoid using the continuous measure. Instead, we classify industries into three distinct categories: those with increasing returns to scale, those with decreasing returns to scale, and the remaining group of unclassified industries. If an industry is identified as featuring IRS in both Bartelme et al. (2024) [7] and Lashkaripour and Lugovskyy (2023) [27], we categorize it as such. Similarly, industries that are consistently classified as DRS in both studies are treated as DRS. For the remaining industries with inconsistent classification between the two sets of estimates, we assign them to a category with undetermined economy to scale and treat them as a reference group in our empirical analysis.

3.3 Testing Proposition 1(iii): FTA Size, Economies of Scale, and Tariff Complementarity

Proposition 1(iii) states that the tariff complementarity effect becomes stronger both with the degree of economies of scale and with the size of an FTA. Accordingly, the reduction in MFN tariffs in response to a 1% preferential tariff cut should be larger when the industry is characterized by increasing returns to scale and the FTA is larger. This prediction motivates the following extension of model (7), which tests whether industries featuring stronger economies of scale experience greater reductions in MFN tariffs when participating in larger FTAs.

$$\Delta_T MFN_{cin} = \sum_{X=I,D} \beta_T^X D_i^X \Delta_T PM_{cin} \times PS_{cin} + \beta_T \Delta_T PM_{cin} \times PS_{cin} + \mu_c + \mu_i + \mu_n + \varepsilon_{cin} \quad (8)$$

Here, D_i^X is an indicator variable that takes the value of 1 if industry i has either IRS ($X = I$) or DRS ($X = D$), and zero otherwise. The unclassified industries serve as the reference group, so the coefficients β_T measure the average tariff complementarity in sectors where the degree of economies of scale is indeterminate. Consistent with Proposition 1(i), we expect $\beta_T > 0$ even for industries in the reference group, reflecting that tariff complementarity holds across all sectors. The main coefficients of interest are β_T^I and β_T^D, which measure how the strength of tariff complementarity differs between IRS and DRS industries relative to this benchmark. According to Proposition 1(iii), ttariff complementarity should be stronger when both the import share of FTA partner countries and the degree of economies of scale are larger; therefore, we expect $\beta_T^I > 0$. Conversely, in industries characterized by diseconomies of scale, the responsiveness of MFN tariffs to preferential tariff cuts should be weaker, implying $\beta_T^D < 0$. Nevertheless, tariff complementarity remains positive even in DRS industries, which requires that $\left(\beta_T + \beta_T^D\right) > 0$.

Column (1) of Table 1 present the benchmark results for tariff complementarity across all industries based on equation (7). As expected, the coefficients β_T are positive and statistically significant for all $T \leq 6$, and the results are very consistent across specifications. This supports prior findings of numerous studies that demonstrate strong evidence of tariff complementarity between internal and external tariffs within an FTA. To contextualize the magnitudes of our estimates, the average trade share across all FTA partners and industries in our sample is 3.8%. Thus, the estimate $\widehat{\beta}_{T=1} = 0.289$ implies that a one percentage point increase in FTA preference margin results in an average reduction in MFN tariff by 0.289×0.038, or by 0.01 percentage points.

This elasticity of MFN tariff with respect to preferential tariff of 0.01 is notably smaller than most estimates in the literature. For example, Estevadeordal et al. (2008) [16] estimate a one-year elasticity of 0.1 for FTAs between a group of Latin American countries, while Mai and Stoyanov (2015) [30] place it in the range of $0.06 - 0.3$ for the Canada-US FTA. The difference between previous estimates and ours primarily arises from the fact that previous studies have largely focused on FTAs between major trade partners with significant bilateral trade shares. In contrast, our results are derived from a broader sample of FTAs, most of which involve smaller countries with relatively modest trade volumes. When applying our estimates to the Canada-US FTA, we find that larger trade volumes result in elasticity estimates of 0.07 for the US and 0.2 for Canada, which are broadly in line with the existing studies.

Next, we turn to the results from specification (8), which estimates tariff complementarity separately for the three industry groups. Our theoretical model predicts that tariff complementarity should hold across all industries, as illustrated in Figure 1, and this prediction is supported by the data. The estimates of β_T for industries with unclassified economy of scale, reported in column (2) of Table 1, are all positive and statistically significant, with magnitudes similar to those derived from the full sample. Similarly, tariff complementarity is present in IRS and DRS industries as well. The estimate of $\left(\beta_T + \beta_T^I\right)$ for IRS and $\left(\beta_T + \beta_T^D\right)$ for DRS industries, shown in columns (5) and (6), are also positive and significant in all years except for one. This confirms that while tariff complementarity still holds even in DRS industries, reinforcing the robustness of the complementarity effect across different industry groups.

More importantly, in line with our theoretical model, we find that tariff complementarity is significantly more pronounced in industries with IRS compared to those with DRS. The estimates of β_T^I are positive and statistically significant across all time horizons, indicating that a one percentage point increase in the preferential tariff margin leads to deeper MFN tariff reductions in industries with IRS relative to the reference group. In contrast, the estimates of β_T^D show no clear pattern, with most of the coefficients being statistically insignificant and indistinguishable from zero. The only statistically significant coefficient among them is negative. Consequently, as indicated in column (7), the difference in tariff complementarity between IRS and DRS industries, $\left(\beta_T^I - \beta_T^D\right)$, is positive for all T and statistically significant at the 10% level in all but two years. Figure 2 presents the coefficients β_T^I and β_T^D from columns (3) and

Table 1. Estimation results for model (E2)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	β_T	β_T^I	β_T^D	$\beta_T + \beta_T^I$	$\beta_T + \beta_T^D$	$\beta_T^I - \beta_T^D$
PM(T=1) × PS	0.289***	0.221***	0.987**	0.009	1.207***	0.230***	0.977*
	(0.053)	(0.050)	(0.502)	(0.093)	(0.501)	(0.079)	(0.506)
PM(T=2) × PS	0.340***	0.312***	0.315*	0.004	0.626***	0.316***	0.311
	(0.043)	(0.049)	(0.191)	(0.074)	(0.189)	(0.064)	(0.196)
PM(T=3) × PS	0.189***	0.175***	0.333***	-0.033	0.508***	0.143***	0.365***
	(0.030)	(0.041)	(0.128)	(0.048)	(0.124)	(0.033)	(0.126)
PM(T=4) × PS	0.376***	0.529***	0.569**	-0.267***	1.098***	0.261***	0.836***
	(0.089)	(0.078)	(0.259)	(0.094)	(0.255)	(0.062)	(0.259)
PM(T=5) × PS	0.337***	0.342***	0.617***	-0.329	0.958***	0.012	0.946***
	(0.024)	(0.014)	(0.133)	(0.233)	(0.134)	(0.234)	(0.256)
PM(T=6) × PS	0.249***	0.218***	0.160**	0.067	0.378***	0.284***	0.093
	(0.029)	(0.011)	(0.068)	(0.059)	(0.068)	(0.059)	(0.087)

Notes: The dependent variable represents the cumulative change in the MFN tariff over the first T years of an FTA. * significant at 10%, ** significant at 5%, *** significant at 1%. Standard errors, shown in parentheses, are calculated at the country levels. All specifications include country and HS6 product fixed effects.

(4), illustrating the dynamic response of MFN tariff changes to FTAs across IRS and DRS industries, with a striking contrast between the two.[8]

The contrast between IRS and DRS industries extends beyond their differential responses and becomes even more striking when we examine the magnitudes of tariff complementarity. The elasticity of MFN tariffs with respect to the FTA preference margin, averaged across all $T \in [1; 6]$, is significantly higher for industries with IRS, standing at 0.031. By comparison, the elasticity in DRS industries is four times smaller than in IRS industries, at just 0.008, while the reference group of unclassified industries has an elasticity of 0.011. This substantial disparity in trade policy responses to FTAs demonstrates the importance of recognizing distinctions between industries based on their economies of scale for an endogenous trade policy analysis. Our results reveal that the strength of tariff complementarity differs significantly across industries with varying economies of scale. Notably, most of the tariff complementarity effect is concentrated in IRS industries, where the impact is much stronger, aligning closely with our theoretical predictions about the role of economies of scale in shaping the response of external tariffs to FTAs.

[8]We report an additional robustness check on pre-FTA MFN tariff changes in Appendix Section 9.4 (Table A1), which shows no evidence of differential pre-trends between IRS and DRS industries.

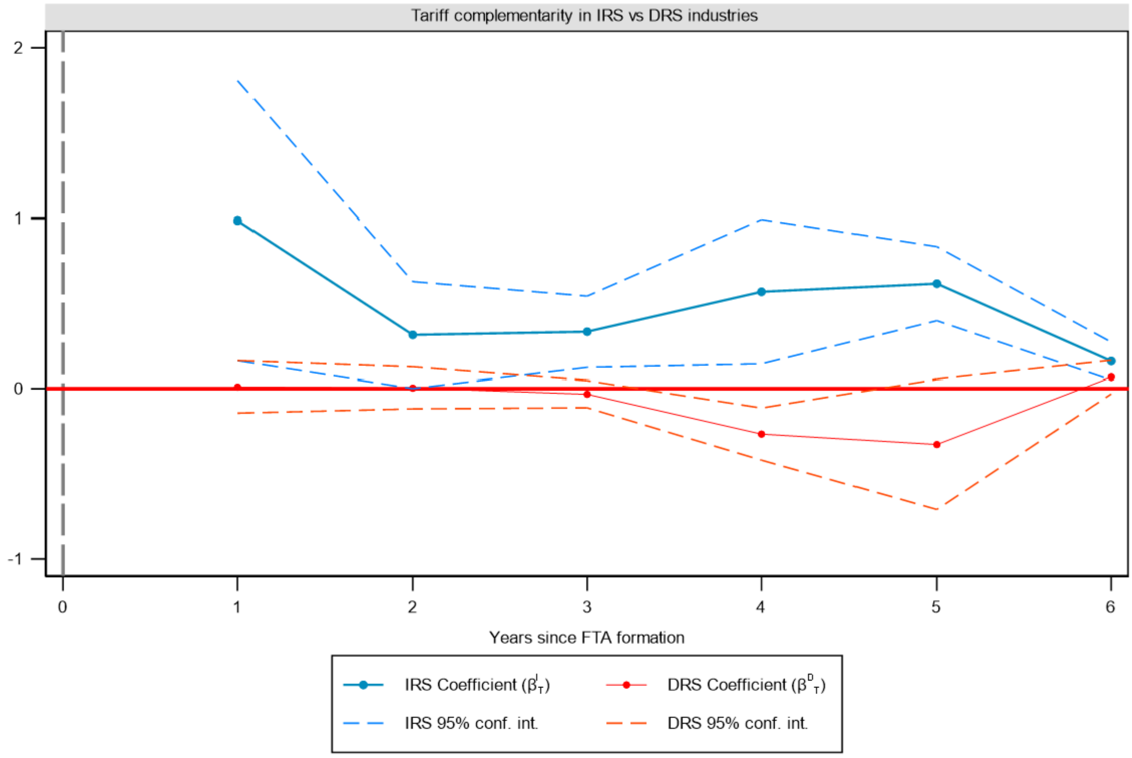


Figure 2: Tariff Complementarity Effects Across IRS and DRS Industries: Estimates from Columns (3) and (4) of Table 1

4 Testing Proposition 1(ii): Tariff Complementarity under the Unrestricted Empirical Specification

In Section 3.3, we introduced the baseline specification (8) which provides a parsimoneous framework for testing part (ii) of Proposition 1. Its appeal lies in simplicity: the test centers on two coefficients, β_T^I and β_T^D, whose predicted signs are derived directly from Proposition 1, part (iii). Evaluating their magnitude and significance provides a straightforward way to assess whether tariff complementarity strengthens with economies of scale, preferential tariff reductions, and preferential market shares.

The drawback of specification (8) is that it is restrictive: most double-interaction and single-variable terms are constrained to zero. To relax these restrictions, we now estimate a fully flexible specification that includes the complete set of interactions and their components:

$$
\begin{aligned}
\Delta_T MFN_{cin} = & \sum_{X=I,D} \beta_T^X D_i^X \times \Delta_T PM_{cin} \times PS_{cin} + \beta_T^0 PS_{cin} \times \Delta_T PM_{cin} + \\
& + \sum_{X=I,D} \gamma_T^X D_i^X \times PS_{cin} + \gamma_T^0 PS_{cin} + \\
& + \sum_{X=I,D} \delta_T^X D_i^X \times \Delta_T PM_{cin} + \delta_T^0 \Delta_T PM_{cin} + \mu_c + \mu_i + \mu_n + \varepsilon_{cin} \qquad (9)
\end{aligned}
$$

Industry dummies D_i^I and D_i^D do not enter individually, as their effects are absorbed by the industry fixed effects μ_i.

Under this extended specification, the difference in tariff complementarity between industries with increasing and decreasing returns to scale can no longer be captured by a single coefficient, as in (8). Instead, the relevant comparisons requires evaluating multiple terms simultaneously. To illustrate, the difference in tariff complementarity between industries in the IRS and DRS groups takes the form:

$$
\begin{aligned}
\Delta_T MFN_{cin} \cdot D_i^I - \Delta_T MFN_{cin} \cdot D_i^D = & \left(\beta_T^I - \beta_T^D\right) \Delta_T PM_{cin} \times PS_{cin} + \\
& + \left(\gamma_T^I - \gamma_T^D\right) PS_{cin} + \left(\delta_T^I - \delta_T^D\right) \Delta_T PM_{cin} \qquad (10)
\end{aligned}
$$

Unlike in the restricted model, where this difference was determined entirely by a single regression coefficient and therefore had a constant sign for any values of $\Delta_T PM$ and PS, expression (10) shows that in the full model the sign and magnitude of the difference may vary with the levels of preferential tariffs and market shares.

To ground this specification in the theoretical framework, we turn to Proposition 1, part (ii). The model predicts that, for any given preferential tariff and market share, tariff complementarity in the IRS industries should be stronger than in the DRS industries. Translating this prediction into empirical terms implies that expression (10) should be positive across the relevant support of $\Delta_T PM$ and PS. To test this, we estimate the full regression model (9) at different time horizons T, substitute the estimated coefficients into (10), and examine whether the implied difference is consistently positive.

We present the results in Figure 3. Because the test derived from equation (10) depends jointly on the values of preferential tariffs and preferential market shares, we place these variables on the axes of the graph. The horizontal axis represents the preferential market share PS, while the vertical axis represents the preferential margin $\Delta_T PM$. To ensure comparability across industries, the scale of each axis is set from zero to the approximately 95^{th} percentile of the distribution of the corresponding variable in the data (16 percentage points for the preferential tariff margin and 0.25 for the preferential market share). The figure contains eight panels, each corresponding to one time horizon T, allowing us to visualize the results across different lags.

Each panel of the figure contains two threshold lines. The solid red line represents the theoretical threshold condition implied by equation (10) when coefficients β-s, γ-s and δ-s are set to the values estimated from equation (9). The dark blue region below this line represents the combinations of $\Delta_T PM$ and PS for which expression (10) is negative, implying that tariff complementarity is stronger in DRS industries than in IRS industries – contrary to the theoretical prediction. Figure 3 shows that such violations are nearly absent across the relevant range of the data. Only for $T = 3$ do we observe a small region where the expression in (10) is negative, and even then it contains only 0.32% of all observations in our data. For other periods, there is even more support for the theoretical prediction that tariff complementarity is greater in IRS industries than in DRS industries, as the regions where condition (10) is negative are very small.

Furthermore, the share of observations in our data that fall into those regions is smaller still: 0.1% in Year 1, 0.06% in Year 2, and 0.02% in Year 8. This means that in the observed data, when we restrict attention to positive preferential tariffs, the fraction of observations for which tariff complementarity would appear stronger in the DRS industries than in the IRS industries is vanishingly small.

The dotted line in Figure 3 provides a conservative check on the robustness of these results. Whereas the solid line is derived directly from condition (10) using the point estimates of the β, γ, and δ coefficients, the dotted line shows what the rejection region would look like if we pushed the parameters to the limits of their statistical confidence. Specifically, we replace these parameters by the values at the bounds of their 95% confidence intervals in a way that maximizes the rejection region: taking the upper bound for $\left(\gamma_T^I - \gamma_T^D\right)$ and the lower bounds for $\left(\beta_T^I - \beta_T^D\right)$ and $\left(\delta_T^I - \delta_T^D\right)$. This combination steepens the threshold curve as much as possible, thereby creating the maximum potential rejection region, highlighted in light blue in Figure 4.

Figure Figure 3 demonstrates that even under this conservative approach, the rejection

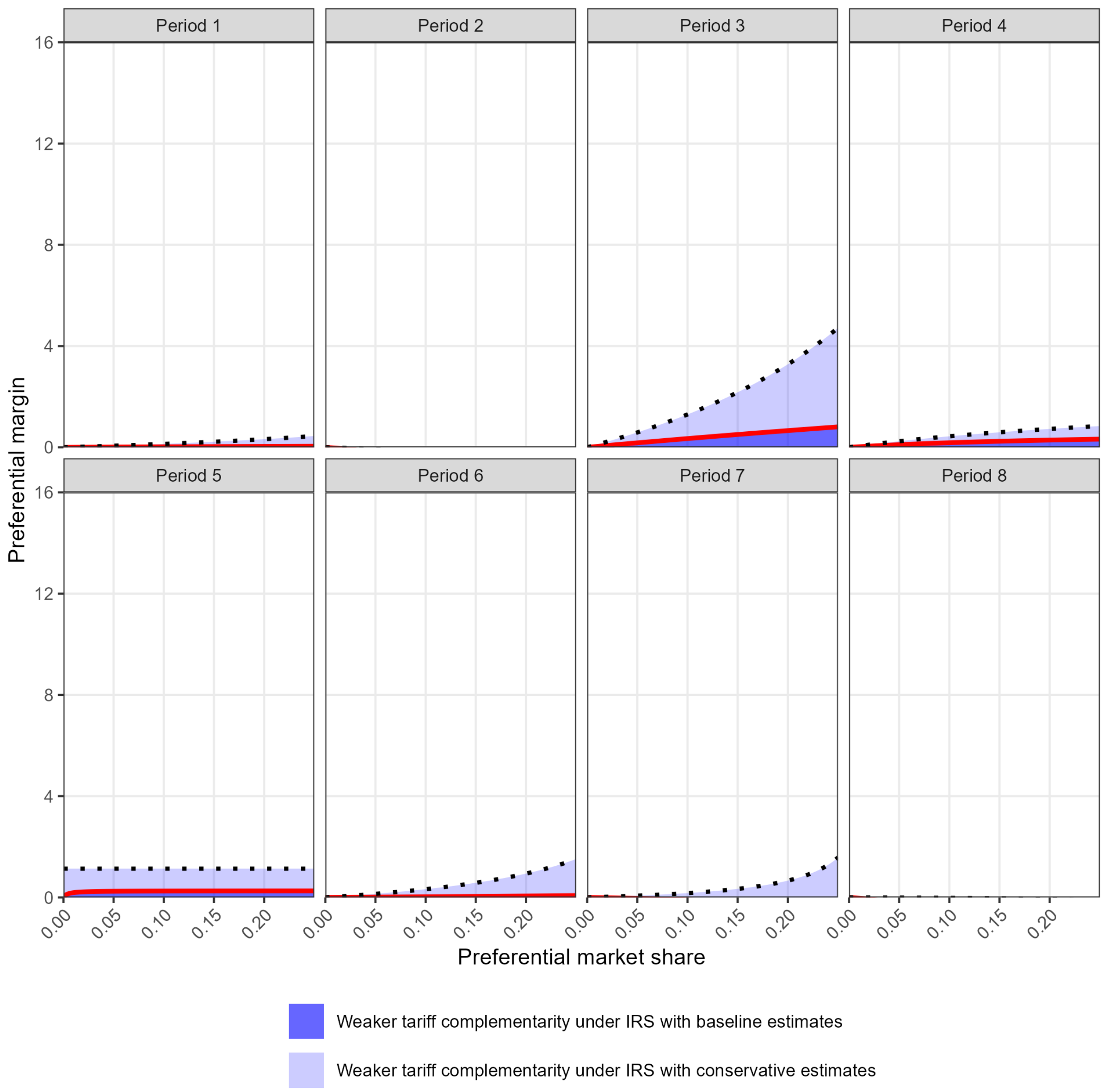


Figure 3: Difference in Tariff Complementarity Effects between IRS and DRS Industries

region remains relatively small. Only for $T = 3$ and $T = 5$ do we observe a visible shaded region, and even then its size is modest relative to the overall support of the variables. Moreover, if the three parameters were independent, the probability of simultaneously drawing values that generate such a region would be on the order of 0.05^3, which is extremelly small. Thus, we view these results as further evidence that tariff complementarity is systematically stronger in IRS industries. Importantly, this conclusion is based on estimates from a fully flexible, unrestricted empirical specification, confirming that our main findings are not driven by the constraints of the baseline model but hold even under the most general empirical framework.

Having both theoretically and empirically established the fact that the external tariffs of member countries are largely affected by different degrees of economies of scale, we next turn our attention to the incentives of countries to form FTAs and how different degrees of economies of scale affect such incentives.

5 Dynamic FTA Formation

In this section, we employ the same oligopoly model of trade outlined in the previous sections and examine the dynamic FTA formation in the WTO world where member countries eliminate their internal tariffs consistent with the Article XXIV of the GATT. We can treat this analysis as the long run where phase-out period ended. We consider three symmetric large open countries i, j, and k. Borrowing the terminology from the existing network literature, countries are treated as nodes in a "graph," and the bilateral FTA links between countries are viewed as edges on this graph.[9] The graph, symbolized as g, represents the array of bilateral FTA connections between countries. Henceforth, g will be referred to as the network of FTAs. Figure 4 provides visual representations of these potential networks. The "empty network" g_ϕ emerges in the absence of any FTAs while g_{ij} is the "FTA insider-outsider network" in which the FTA non-member k is the "FTA outsider" and the FTA members i and j are "FTA insiders". g_i^H is the "hub-spoke" network where countries j and k are "spokes" and each have an FTA with country i who is the "hub". Finally, g^F is the "global free trade network" where all countries are linked through FTAs.

Under the empty network g_ϕ, Article I of the GATT forbids tariff discrimination and thus all countries simultaneously choose their respective non-discriminatory MFN tariffs: $t_{ij}(g_\phi) = t_{ik}(g_\phi) = t_i(g_\phi) \equiv \arg\max w_i(g_\varnothing)$. In the case of an FTA insider-outsider network g_{ij}, unlike our previous tariff analysis, internal tariffs are eliminated between insiders: $t_{ij} = t_{ik} = 0$ while each country retains autonomy in setting its owns external tariffs on the outsider country. The outsider country, in turn, imposes the MFN tariff on both insiders. Similarly, under a hub-spoke network g_i^H, the hub country imposes zero external tariffs on spokes while spokes impose external tariffs on each other: $t_{ij} = 0, t_{ji} = 0, t_{ik} = 0, t_{ki} = 0$.

In this section, we focus on optimal applied tariffs. Our first finding suggests that, regardless of the FTA networks, countries have incentives to impose higher external tariffs as δ rises,

[9] Some notable works that have used this terminology for PTAs include Goyal and Joshi (2006) [19], Furusawa and Konishi (2007) [18], Xue et al. (2013) [50], Lake and Yildiz (2016) [24], Lake (2017) [23], Lake et al. (2020) [26], and Nken and Yildiz (2022) [35].

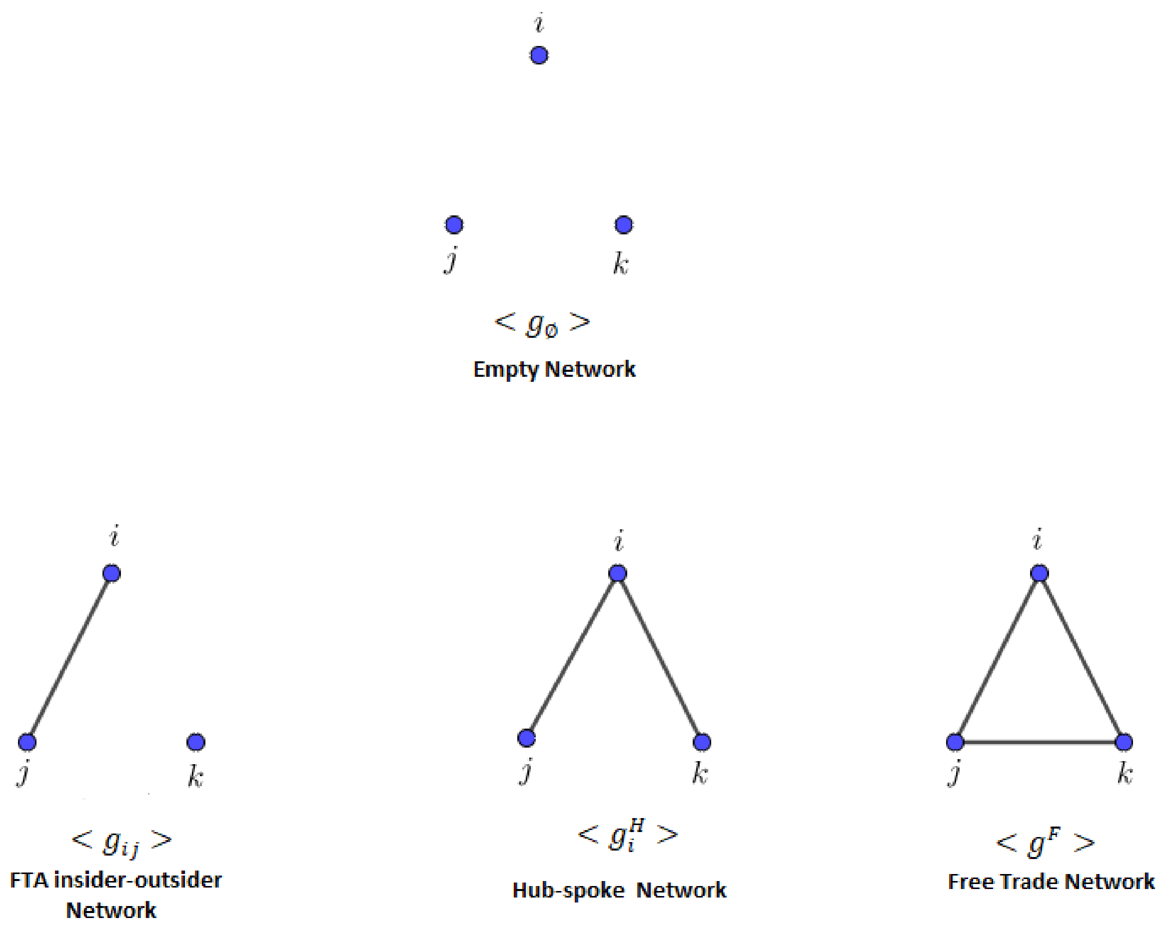


Figure 4: Network Position

implying higher optimal tariffs with a greater degree of economies of scale:

$$\frac{\partial t_i(g)}{\partial \delta} > 0, \text{ where } g = g_\phi, g_{ij}, g_{ik}, g_{jk}, g_j^H \text{ and } g_k^H \tag{11}$$

The intuition behind the above result is as follows: as δ rises, an increase in the production leads to lower unit cost of production leading to a higher mark. As a result, each country has incentive to impose higher external tariffs to protect domestic market and extract rent from exporters.

First, we confirm our result in Proposition 1 regarding the tariff complementarity: regardless of the degree of economies of scale, upon formation of a bilateral FTA, insiders have an incentive to lower their tariffs on the FTA outsider and this tariff complementarity effect deepens as the degree of economies of scale becomes stronger: $\Delta t_{ik}(g_\phi - g_{ij}) = t_i(g_\phi) - t_{ik}(g_{ij}) > 0$ for all δ and $\frac{\partial \Delta t_{ik}(g_\phi - g_{ij})}{\partial \delta} > 0$. Then, we find that the same holds for an additional bilateral FTA under g_{ij}, leading to a hub and spoke regime: $\Delta t_{kj}(g_{ij} - g_i^H) = t_{kj}(g_{ij}) - t_{ik}(g_i^H) > 0$

for all δ and $\frac{\partial \Delta t_{kj}(g_{ij}-g_i^H)}{\partial \delta} > 0$.

As indicated before, the existence of economies 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. This implies that FTA formation affect the external tariffs of the non-member country under a bilateral FTA. Next, we turn our attention to how a bilateral FTA formation affects the outsider's tariffs. We know from Missios et al. (2016) [32] and Saggi et al. (2018) [46] that the FTA outsider has an incentive to reduce its tariffs on the imports from FTA insiders due to external trade diversion under perfectly competitive markets with increasing marginal costs. The intuition is as follows. When two countries liberalize trade on preferential basis, they import more from each other and start exporting less to the third country. This phenomenon is called as external trade diversion, leading higher elasticity of export supply of insiders and thus lower external tariff of the outsider. Unlike the perfectly competitive markets in Missios et al. (2016) [32] and Saggi et al. (2018) [46], we employ an oligopoly model of trade with flexible degrees of economies of scale and argue that how a bilateral FTA formation changes the external trade diversion and thus the response of the FTA outsider's external tariff heavily depend upon the nature of the economies of scale. We find that, as in Missios et al. (2016) [32] and Saggi et al. (2018) [46], external trade diversion arises and a bilateral FTA formation induces the FTA outsider to lower its tariffs on imports from FTA insiders only when there is diseconomies of scale: $\Delta t_k(g_\phi - g_{ij}) = t_k(g_\phi) - t_k(g_{ij}) > 0$ if and only if $\delta < 0$. We know from the existing literature that, when there exists constant returns to scale, market segmentation arises and thus markets are strategically separated. Under such a case, FTA formation has no impact of the external tariff of the outsider relative to the empty network. However, when the industry exhibits economies of scale ($\delta > 0$), larger production of FTA insiders due to preferential access via FTA formation leads to lower average cost of production and higher mark-up for insiders. As a result, the outsider has a greater incentive to extract rent from insiders and thus raise its tariff on imports from insiders relative to g_ϕ. The opposite result obtains when the industry exhibits diseconomies of scale. The above results are represented in Figure 5 and lend support to Missios et al. (2016) [32] and Saggi et al. (2018) [46] when there is diseconomies of scale and provides additional insights on the tariff response of the outsiders when industries exhibit economies of scale.

Would the above result extend to the case where there is g_{ij} and i and k form an additional

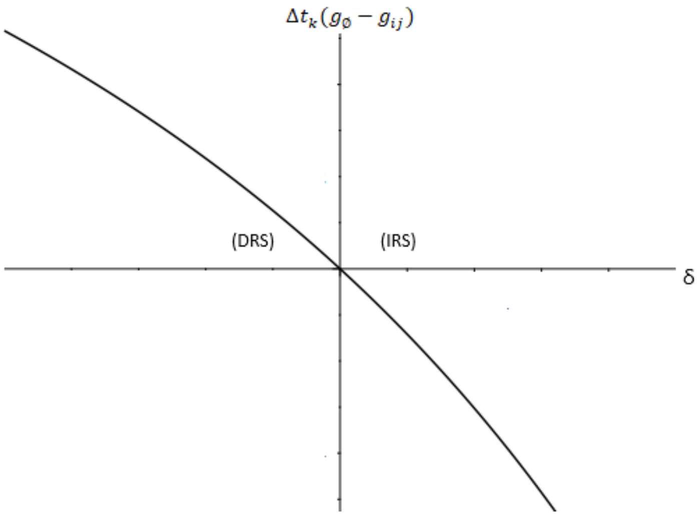


Figure 5: Tariff Complementarity Effect: FTA-outsider

FTA, leading to a deviation from g_{ij} to the hub and spoke regime g_i^H? We argue that it would not be the case and the opposite result obtains. Note first that the outsider country (j) of such an FTA is already a partner of country i and thus can not be identified as a true outsider. Second, since it gives up its unconstrained domestic market protection, country k's total output declines when it forms an FTA with country i and becomes a spoke country under g_i^H. This is due to the fact that the decline in output in the domestic market exceeds the limited increase in the export in the hub's market. This implies that country k's firm's mark up falls (rises) under economies (diseconomies) of scale. As a result, country j's incentive to extract rent from country k's firm weakens (strengthens) under economies (diseconomies) of scale.

Next, we introduce our equilibrium concept and strategies.

6 Equilibrium Theory: Concepts and Strategies.

Building upon the models by Seidmann (2009) [48], Lake and Yildiz (2016) [24] and Lake et al. (2020) [26], we employ a dynamic FTA formation model that initiates from an empty network. Consistent with Ornelas (2008) [39] and Ornelas and Liu (2012) [40], we assume that at most one FTA can be formed in one period and that FTAs formed in previous periods are binding and cannot be severed.[10] We construe a period as the duration necessitated for negotiating an agreement, thereby tying our model more closely to empirical realities in the

[10]The binding nature of trade agreements is pervasive in the literature and realistic in terms of real-world observation. It can represent a reduced form for more structural justifications such as sunk costs (see McLaren (2002) [31] for empirical support).

field.[11] Let g_{t-1} denote the network at the beginning of time t, and the transition of network in the current period be denoted by $g_{t-1} \to g_t$. As an example, we denote the transition of networks $g_{t-1} \to g_t \to g_{t+1}$, where the network g_{t-1} evolves into g_{t+1} through g_t and then remains at g_{t+1} forever. Finally, our network remains permanent when one of two conditions occur: either no agreement forms in a given period or once global free trade emerges.[12] First, when no agreement forms in a given period, the assumption of Markov strategies implies no agreement forms in any subsequent periods. Second, once global free trade emerges, the assumption that previously formed agreements cannot be severed implies global free trade remains forever. Table 2 depicts the feasible transitions for the current period, moving from g_{t-1} to g_t, where g_t represents the network at the end of the current period. A path of networks, $g_{t-1} \to g_t \to g_{t+1}$, indicates a progression beginning at g_{t-1}, passing through g_t, and ultimately settling permanently at g_{t+1}. Consistent with our model, the network remains unchanged from the third period onwards. Alternatively, the path of networks $g_\phi \to g_\phi$ denotes a situation where no FTA formation takes place.

Countries hold preferences over various paths of FTA networks, which is illustrated by the continuation payoffs. Let β denote the discount factor and $V_i(g_t)$ denote the continuation payoff from the path of networks. Assuming that the network at the beginning of the current period is empty, denoted by g_ϕ, and a country i's one-period payoff is equivalent to the national welfare $W_i(g)$, the continuation payoff for country i from the path of networks $g_\phi \to g_{ij} \to g_i^H \to g^F$ is defined as $V_i(g_{ij}) = W_i(g_{ij}) + \beta W_i(g_i^H) + \frac{\beta^2}{1-\beta} W_i(g^F)$. In a different scenario, if no FTA forms in the current period, the empty network g_ϕ will remain indefinitely. Hence, the continuation payoff for country i from the path of networks $g_\phi \to g_\phi$ will be $V_i(g_\phi) = \frac{1}{1-\beta} W_i(g_\phi)$.

We employ a deterministic protocol as in Seidmann (2009) [48], Lake and Yildiz (2016) [24], Lake et al. (2020) [26], where the "leader" country has the first opportunity in each period to propose an FTA that forms if and only if recipient countries accept the proposal. If the leader country does not have a proposal accepted by the follower countries, then one of the follower countries has the opportunity to propose an FTA.[13]

[11] This interpretation aligns with the observation by Odell (2006, p. 193) [36] that NAFTA negotiations began in 1986 but the agreement wasn't signed until 1992. Similarly, Mölders (2012, 2015) [33, 34] and Freund and McDaniel (2016) [17] provide additional empirical evidence, illustrating the extended time frames often associated with FTA negotiations.

[12] These conditions are expected to take place no later than the third period, a feature that is vital in applying backward induction within our model.

[13] Unlike Aghion et al. (2007) [1] where only the leader country can make proposals, our protocol empowers the follower countries to propose FTAs and thus follower countries can form their own FTA.

Without loss of generality, from hereon, we assume that country a is the leader country and country b and c are the recipients. Due to complete symmetry, multiple equilibria could arise. In order to eliminate this possibility, we assume some infinitesimal non-economic benefit $\epsilon > 0$ that countries b and c obtain when forming a PTA with country a. In other words, country a is assumed to be the the leader country that is the most attractive PTA partner. Similarly, country a receives some $\epsilon > 0$ when forming with country b rather than country c. Formally, stages 1-2 describe the protocol in every period:

- **Stage 1(a):** Country a proposes an FTA to a country. If the recipient country accepts the proposal, then the period ends. If the recipient country rejects the proposal, then the game moves to stage 1(b). Otherwise, the game moves to stage 2.
- **Stage 1(b):** Country a proposes to the country that didn't reject its proposal in stage 1(a). If FTA forms then the period ends. Otherwise, the game moves to stage 2.
- **Stage 2:** Country b proposes an FTA, and the period ends no matter what happens in this stage.

The series of the above stages indicate that the network remains unchanged once global free trade is established or when a period ends without any FTA formation. Table 3 represents the available proposals for each country acting as the proposer. The set of proposals for country i is denoted as $S_i(g)$, where $\rho_i(g) \in S_i(g)$ denotes a propsal. Once a proposal $\rho_i(g)$ from country i is received, every recipient country j, part of the proposed agreement, offers a response $r_j(g, \rho_i(g)) \in \{Y, N\}$, where '$Y$' signifies acceptance and 'N' implies rejection. For each subgame within network g, the Markov strategy of country i should accomplish two primary tasks: (i) designate a proposal $\rho_i(g) \in S_i(g)$ for the stage where it serves as the proposer, and (ii) allocate a response $r_i(g, \rho_j(g)) \in \{Y, N\}$ to any potential proposal it might receive from another country j. We aim to find a pure strategy Markov perfect equilibrium.[14] Specifically, using backward induction, we find a pure strategy subgame perfect equilibrium where the current period's proposal by the proposer and the responses by the respondent depend only on the history through the network established at the end of the preceding period.

[14] Our approach aligns with the method outlined by Seidmann (2009) [48].

Network at the beginning of Current Period	Network at the end of Current Period
g_ϕ	g_ϕ, g_{ij}, g_{ik}, g_{jk}
g_{ij}	g_{ij}, g_i^H, g_j^H
g_i^H	g_i^H, g^F
g^F	g^F

Table 2: Pathways of Feasible Networks Transitions

	$S_i(g)$	$S_j(g)$	$S_k(g)$
g_ϕ	$\{\phi$, ij, ik $\}$	$\{\phi$, ij, jk$\}$	$\{\phi$, ik , jk$\}$
g_{ij}	$\{\phi$, ik$\}$	$\{\phi$, jk$\}$	$\{\phi$, ik , jk$\}$
g_i^H	$\{\phi\}$	$\{\phi$, jk$\}$	$\{\phi$, jk$\}$
g^F	$\{\phi\}$	$\{\phi\}$	$\{\phi\}$

Table 3: Action Space of Proposer Countries Across FTA Subgames

7 Equilibrium path

In this section, we examine the driving forces behind the FTA formation under optimal tariffs. In a dynamic model of FTA formation, the tension between myopic and farsighted incentives of the FTA outsider and FTA insiders creates the possibility of dynamic free riding and dynamic exclusion incentives. As a result, the equilibrium path of networks revolve around the dynamic free riding incentive of an FTA outsider and the dynamic exclusion incentive of FTA insiders. These dynamic incentives heavily depend upon the degree of economies of scale and how patient the countries are.

7.1 Myopic and Farsighted incentives

Let $w_i(g)$ denote country i's welfare as a function of network g and $\Delta w_i(g' - g)$ represent the difference between country i's welfare under networks g_0 and g, where $g' = g + ij$ denotes the trade network that results when an FTA between countries i and j is added to network g. In our dynamic model, incentives that involve payoff comparisons in the current period and in the immediate next period are called *myopic incentives* that capture mainly short-run considerations. When governments make payoff comparisons between the current period and periods beyond the immediate next period, we call these incentives as *farsighted incentives* that involve long-run considerations. Specifically, what are the myopic and farsighted considerations of an FTA outsider under a bilateral FTA? On one hand, myopically, an FTA outsider decides whether to become a spoke country under a hub-and-spoke agreement relative to being an outsider under a bilateral FTA. On the other hand, farsightedly, the decision of an FTA outsider is between becoming a spoke on the path to global free trade and staying as an outsider. Similarly, an FTA insider myopically decides whether to become a hub country under a hub-and-spoke agreement relative to being an insider under a bilateral FTA while farsightedly, the decision of an FTA insider is between becoming hub on the path to global free trade and staying as an insider. Lemma 1 describes the myopic incentives that are pivotal for the equilibrium path of FTA networks.

Lemma 1. *The following myopic incentives obtain for all degrees of economies of scale:*

(i) ***(Participation incentive):*** *Both insiders and the outsider benefit from a bilateral FTA formation relative to an empty network, and FTA insiders are better off relative to being an FTA outsider:* $\Delta w_i(g_{ij} - g_\phi) > \Delta w_i(g_{jk} - g_\phi) > 0$ *for all* δ *and the incentive to form a bilateral FTA rises with the degree of economies of scale:* $\frac{\partial \Delta w_i(g_{ij} - g_\phi)}{\partial \delta} > 0$.

(ii) ***(Spoke-spoke FTA):*** *Two spoke countries always have incentives to form the final FTA:* $\Delta w_j(g^F - g_i^H) > 0$ *for all* δ*, and such incentives get stronger with the degree of economies of scale:* $\frac{\partial \Delta w_j(g^F - g_i^H)}{\partial \delta} > 0$.

(iii) ***(Myopic incentive to become hub):*** *An FTA insider always has an incentive to form an additional FTA to become the hub:* $\Delta w_i(g_i^H - g_{ij}) > 0$ *for all* δ*, and* $\frac{\partial \Delta w_i(g_i^H - g_{ij})}{\partial \delta} > 0$ *holds.*

(iv) ***(Myopic Free Riding Incentive):*** *An FTA outsider has no incentive to become a spoke:* $\Delta w_k(g_{ij} - g_i^H) > 0$ *for all* δ*, and* $\frac{\partial \Delta w_k(g_{ij} - g_i^H)}{\partial \delta} > 0$.

The first three parts of the above lemma govern the attractiveness of a bilateral FTA formation. The first part of the lemma informs us that a bilateral FTA formation is Pareto improving relative to empty network while FTA insiders benefit more relative to the outsider. Intuitively, the mutual exchange of preferential access in each others' market benefits the insiders and such benefit and the external tariff insiders face in the outsider's market relative to empty network heavily depend on the degree of economies of scale. Consider first the scenario where industry exhibits constant returns to scale. Under such a case, outsider's tariff on insiders remains unchanged and insiders benefit from the bilateral FTA formation purely due to exchange of preferential access in each others markets. While the outsider faces discrimination in the insiders' markets, it faces a lower external tariffs due to tariff complementarity while keeping its optimum external tariffs in the domestic market. Therefore, the outsider also benefits from the FTA formation while the insiders' benefit from preferential access is larger than the outsider's benefit from the tariff complementarity. Suppose now that the industry exhibits economies of scale. Under such a case, insiders produce larger output due to preferential access, leading to extra benefits from lower average costs. From our tariff discussion, we know that insiders lower their external tariffs on outsiders while facing higher tariffs in the outsider's market relative to no agreement. Our finding suggests that the extra benefit due to economies of scale dominates the tariff effect in such a way that the welfare gain of insiders from the bilateral FTA formation rises with the degree of economies of scale: $\frac{\partial \Delta w_i(g_{ij}-g_{\phi})}{\partial \delta} > 0$. Here one should note that, relative to no agreement, the outsider's output rises even more than the insiders due to lower tariffs in the insider's markets and higher tariffs in the domestic market and thus the outsider also benefits from the bilateral FTA formation but not as much as the insiders due to tariff distortions. Finally, consider the scenario where the industry exhibits diseconomies of scale under which insiders enjoy deeper preferential access in each other's markets and lower tariffs in the outsider's market. However, larger output implies higher average cost. Our findings suggest that the tariff effects dominates the cost effect and the insiders still gain from the bilateral FTA formation. On the flip side, the outsider now faces higher tariffs in the insiders' markets and reduces its tariffs on the insiders. Moreover, the output of the outsider is still larger under a bilateral FTA relative to no agreement, leading to a higher average cost.

Part (ii) of the above lemma informs us that, since spoke countries under a hub-and-spoke regime suffer from the discrimination they face against each other, they always have incentives to form the final FTA leading to global free trade for all δ and this incentive gets stronger when the industry exhibits greater degree of economies of scale: $\frac{\partial \Delta w_j(g^F-g_i^H)}{\partial \delta} > 0$.

We show in Part (iii) that an FTA insider always has an incentive to form another independent FTA to become the hub country since it has sole preferential access in both spoke markets. This implies that an FTA insider always has a myopic incentive to become the hub country and such incentive becomes stronger with higher degree of economies of scale: $\frac{\partial \Delta w_i(g_i^H-g_{ij})}{\partial \delta} > 0$.

Finally, we address the following question: is a bilateral FTA formation always attractive?

Part (iv) of the above lemma informs us that it is not the case when an FTA outsider decides whether to become a spoke under the hub and spoke regime. On one hand, the FTA outsider benefits from lower tariffs faced when exporting to the FTA insiders due to tariff complementarity (deeper when there exist economies of scale). On the other hand, the FTA outsider faces discrimination in the FTA insiders' markets and becoming a spoke eliminates this discrimination in the hub's market. Here, it is important to note that becoming a spoke does not benefit the outsider greatly since the other spoke already has an FTA with the hub country and thus the additional FTA does not come with any preferential access. The tariff complementarity benefit and being able to impose optimal tariffs as an outsider together dominate the benefit of eliminating discrimination in one of the insiders' markets (insider becoming the hub country) and thus we argue that the FTA outsider always has a *myopic free riding incentive* that gets larger as the industry exhibits greater degree of economies of scale: $\Delta w_k(g_{ij} - g_i^H) > 0$ for all δ and $\frac{\partial \Delta w_k(g_{ij} - g_i^H)}{\partial \delta} > 0$.

We next turn our attention to the farsighted incentives that play a crucial role in the formation of FTAs.

Lemma 2. *The following farsighted incentives obtain for all degrees of economies of scale:*

(i) ***(Farsighted incentive to become a spoke):*** *FTA outsider prefers global free trade network to remaining as an outsider under a bilateral FTA:* $\Delta w_i(g^F - g_{jk}) > 0$ *for all* δ *and this preference is stronger as the degree of economies of scale rises:* $\frac{\partial \Delta w_i(g^F - g_{jk})}{\partial \delta} > 0$

(ii) ***(No farsighted exclusion Incentive):*** *an FTA insider always prefers free trade to being an insider under a bilateral FTA:* $\Delta w_i(g^F - g_{ij}) > 0$ *for all* δ *and this preference is stronger as the degree of economies of scale rises:* $\frac{\partial \Delta w_i(g^F - g_{ij})}{\partial \delta} > 0$

It is immediate from part (i) of the above lemma that the FTA outsider always has a farsighted incentive to participate in FTA expansion that ultimately yields global free trade, regardless of the degree of economies of scale. While the FTA outsider always has a myopic free riding incentive and thus has no incentive to eliminate discrimination in one of the FTA insiders' markets via becoming spoke in a hub and spoke regime (see Lemma 1), the first part of the above lemma informs us that it benefits from eliminating the discrimination faced in both FTA insider markets. The intuition behind this result can be explained as follows. Consider initially a bilateral FTA under the scenario where the industry exhibits diseconomies of scale

($\delta < 0$). It is immediate from our previous tariff discussion under an FTA that, relative to no agreement, the non-member country is able to protect its market with lower MFN tariffs on exports from both members while its firm faces a higher discriminatory tariffs in the export markets of members. Our result confirms that, relative to being an outsider under an FTA, the non-member country benefit from having free access in both export markets even when higher output levels lead to higher average costs under diseconomies of scale. As δ rises, non-member country has an incentive to raise its MFN tariff on exports from members while enjoying lower external tariffs in the export markets. However, we should note here that the firm of the non-member country is discriminated against its rival exporter, and thus the benefit from the tariff complementarity in the members' markets is limited. With stronger economies of scale, the improved efficiency effect of freer trade is so dominant that it outweighs the ability to impose higher MFN tariffs and facing lower external tariffs under an FTA. This is a critical result since we show later that the tension between the FTA outsider's myopic free riding incentive and its farsighted incentive to participate in FTA expansion determines the dynamic free riding incentive. Moreover, the degree of economies of scale plays a crucial role in evaluating the strength of the tension between myopic and farsighted incentives of the FTA outsider.

The intuition behind part (ii) of the above lemma unfolds as follows: FTA insiders have no farsighted benefit of permanently excluding the FTA outsider from subsequent FTA formation that delivers global free trade. Ultimately, global free trade benefits the FTA insiders by removing the FTA outsider's tariff barriers. This benefit gets larger as the economies of scale get stronger since larger output leads to more efficiency gains via lower average cost of production: $\frac{\partial \Delta w_i(g^F - g_{ij})}{\partial \delta} > 0$. Comparing a country's continuation payoffs across paths of FTA networks often reveals tensions between myopic and farsighted incentives. However, while the part (iii) of Lemma 1 argues that the myopic incentive to become hub under a hub and spoke regime always exists, part (ii) of the Lemma 1 informs us that FTA insiders do not have a *farsighted exclusion incentive* and thus member countries have an incentives to engage in subsequent FTA formation leading to global free trade. Therefore, for insiders of an FTA, there exists no tension between myopic and farsighted incentives and dynamic exclusion incentives do not arise.

We now solve the equilibrium path of networks in the FTA game by backward induction and argue that the existence of dynamic free riding incentives is crucial..

7.2 Subgames at hub-spoke networks

To begin the backward induction, consider a subgame at a hub-spoke network g_i^H. The following lemma follows directly from Lemma 1(ii).

Lemma 3. *In the subgame at a hub-spoke network* g_i^H*, spoke countries form the final FTA that leads to global free trade:* $g_i^H \to g^{FT}$.

7.3 Subgames at FTA insider-outsider networks

We now roll back to a subgame at an insider-outsider network g_{ij}^{FTA}. As discussed above, since two members have no farsighted incentive to exclude the third country from free trade network and each member always has an incentive to become hub on the path to global free trade, there exists no tension between myopic and farsighted incentives for the member countries and thus the dynamic exclusion incentives never arise. However, the tension between myopic and farsighted FTA formation incentives exist for the FTA outsider leads to dynamic free riding incentives. We next explore this tension facing the FTA outsider.

7.3.1 Dynamic free riding incentive

A tension between myopic and farsighted incentives for the FTA outsider creates the possibility of a dynamic free riding incentive. Part (i) of Lemma 2 argues that, regardless of the degree of economies of scale, the elimination of discrimination in both members' market drives the farsighted incentive of an FTA outsider to participate in FTA expansion that yields global free trade and such incentives rise with the degree of economies of scale: $\Delta w_k(g^F - g_{ij}) > 0$ for all δ and $\frac{\Delta w_k(g^F - g_{ij})}{\partial \delta} > 0$. However, part (iv) of Lemma 1 informs us that an FTA outsider has a myopic free riding incentive, and hence a myopic incentive to refuse subsequent FTA formation and as the degree of economies of scale rises this myopic incentive becomes stronger: $\Delta w_k(g_{ij} - g_i^H) > 0$ for all δ and $\frac{\Delta w_k(g_{ij} - g_i^H)}{\partial \delta} > 0$. Here, the discount factor mediates the tension between the FTA outsider's farsighted incentive to become a spoke on the path to global free trade and its myopic free riding incentive to resist to end up as a spoke country. Specifically, the FTA outsider prefers becoming a spoke rather than remaining a permanent FTA outsider when

$$w_k\left(g_i^H\right)+\frac{\beta}{1-\beta}w_k\left(g^F\right)>\frac{1}{1-\beta}w_k\left(g_{ij}\right) \tag{12}$$

and this holds when FTA outsider is sufficiently patient:

$$\beta>\bar{\beta}_{OUT}(\delta)\equiv\left[1+\frac{\Delta w_k\left(g^F-g_{ij}\right)}{\Delta w_k\left(g_{ij}-g_i^H\right)}\right]^{-1} \tag{13}$$

In other words, sufficiently myopic ($\beta<\bar{\beta}_{OUT}(\delta)$) FTA outsider refuses subsequent FTA formation and in such a case we argue that the FTA outsider has a *dynamic free riding incentive*. When we examine how $\bar{\beta}_{OUT}(\delta)$ changes with the degree of economies of scale, we find that $\bar{\beta}_{OUT}(\delta)$ is a convex function of δ with reaching a minimum point when there is a CRS. Relative to CRS, the extent that the FTA outsider holds a dynamic free riding incentive rises (i.e. $\bar{\beta}_{OUT}(\delta)$ rises) as the degree of economies of scale rises since the myopic free riding incentive $\Delta w_k\left(g_{ij}-g_i^H\right)$ rises at a faster rate relative to the farsighted incentive of becoming a spoke $\Delta w_k\left(g^F-g_{ij}\right)$: $\frac{\partial\left(\frac{\Delta w_k\left(g_{ij}-g_i^H\right)}{\Delta w_k\left(g^F-g_{ij}\right)}\right)}{\partial\delta}>0$. On the other hand, as the degree of diseconomies of scale rises (δ falls and become more negative) relative to CRS, the myopic free riding incentive falls at a slower rate relative to the farsighted incentive of becoming a spoke. As a result, the critical discount factor is at minimum when there is constant returns to scale and rises as δ rises or falls. The Figure 6 illustrates the dynamic free riding incentive and how it varies with the degree of economies of scale. The following proposition summarizes our findings:

Proposition 2. *The equilibrium path of networks in the FTA game is:*

(i) $g_\phi\rightarrow g_{ij}\rightarrow g_i^H\rightarrow g^F$ *when* $\beta>\overline{\beta}_{OUT}(\delta)$

(ii) $g_\phi\rightarrow g_{ij}$ *when* $\beta<\overline{\beta}_{OUT}(\delta)$

(iii) $\frac{\partial^2\overline{\beta}_{OUT}(\delta)}{\partial\delta^2}>0$ *and* $\frac{\partial\bar{\beta}_{OUT}(\delta)}{\partial\delta}=0$ *when* $\delta=0$.

Note that dynamic free riding incentive prevents attaining free trade in the equilibrium path and gets stronger under both economies and diseconomies of scale relative to CRS. This is an important result since most of the literature focusing on whether the formation of FTAs acts as building or stumbling blocs for attaining global free trade assumes that there is a constant returns to scale. Our findings suggest that the equilibrium path of FTA formation is less likely

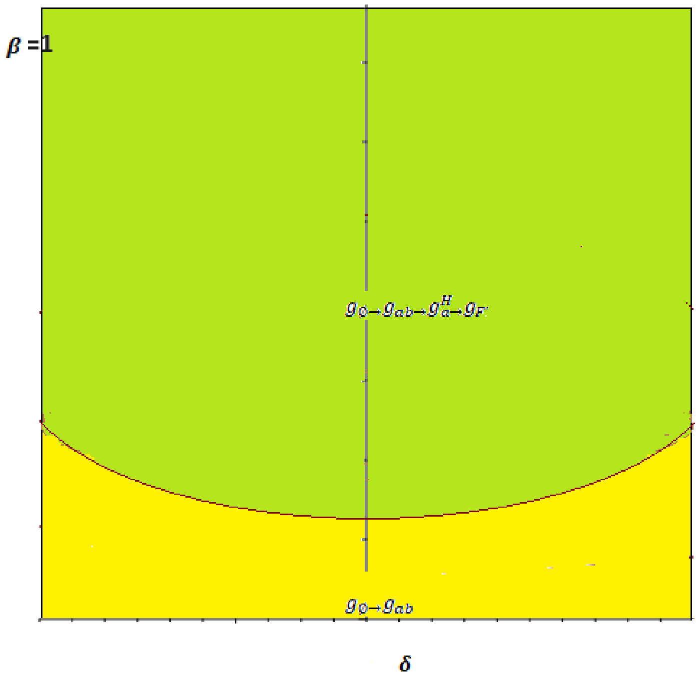


Figure 6: Dynamic Free Riding Incentive

to lead to global free trade under economies and diseconomies of scale relative to CRS when the countries are sufficiently myopic.

8 Conclusion

One of the striking realities of today's global trade environment is the widespread prevalence of free trade agreements. In the rather extensive FTA literature, the role of economies of scale on the external tariff setting and the extent of FTA formation is largely ignored despite the crucial implications of different degrees of economies of scale in determining the trade and trade policy effects of FTAs. The present paper aims at providing a comprehensive theoretical and empirical analysis of the role of economies of scale on the external tariff setting under FTAs. Our model of imperfect competition with flexible degrees of economies of scale yields that FTA member countries have an incentive to reduce their external MFN tariffs following the FTA formation regardless of the degree of economies of scale and this tariff complementarity effect gets stronger as the degree of economies of scale rises. We provide empirical evidences and confirm that tariff adjustments following the FTA formation are consistent with theoretical findings, emphasizing the important role of including the degree of economies of scale in the FTA analysis. Then, focusing on a WTO world with three symmetric countries, we employ a dynamic FTA formation model to examine how the extent of FTA formation is affected by different degree of economies of scale. Our findings informs us that the equilibrium path of FTA formation is fundamentally shaped by the degree of economies of scale in production technology.

9 Appendices

9.1 Cost Functions: one-to-one correspondence

Let Cobb-Douglas production function be defined as: $Q = K^{\alpha}L^{\beta}$ where the degree of economies of scale is captured by $\alpha + \beta$. Then, via cost minimization, we can obtain the corresponding cost function as: $C = AQ^{\alpha+\beta}$ where A is normalized to be 1. We next show that there is a one-to-one correspondence between δ in our cost function and the scale parameter $\alpha + \beta$ in the general form of the Cobb-Douglas cost function. By equating both cost functions we obtain:

$$\alpha + \beta = (\frac{\ln(cQ - \delta Q^2)}{\ln(Q)})^{-1}$$

For any c and Q, Figure A1 shows the one-to-one correspondence between the parameters δ and $\alpha + \beta$.

9.2 Optimal Tariffs under FTAs

Consider $i, p \in M$ and $z, k \in N$. The Consumer Surplus for country i, denoted as CS_i, and its partial derivative with respect to t_i are given as:

$$\frac{dCS_i(f)}{dt_i} = x_i^f \cdot \frac{dx_i^f}{dt_i},$$

where,

$$x_i^f = x_{ii}(f) + x_{ki}(f) + \sum_{z \in M, z \neq i} x_{pi}(f) + \sum_{z \in N, z \neq k} x_{zi}(f)$$

The Tariff Revenue for country i, denoted as TR_i, and its partial derivative with respect to t_i are given as:

$$TR_i(f) = \sum_{z \in M, z \neq i} \tau x_{pi}(f) + t_i(f)x_{ki}(f) + (n - m - 1)t_i(f)x_{zi}(f)$$

$$\frac{d}{dt_i}TR_i(f) = \sum_{z \in M, z \neq i} \tau \frac{dx_{pi}(f)}{dt_i} + \left(x_{ki}(f) + t_i(f)\frac{dx_{ki}(f)}{dt_i}\right) + \sum_{z \in N, z \neq k} \left(x_{zi}(f) + t_i(f)\frac{dx_{zi}(f)}{dt_i}\right)$$

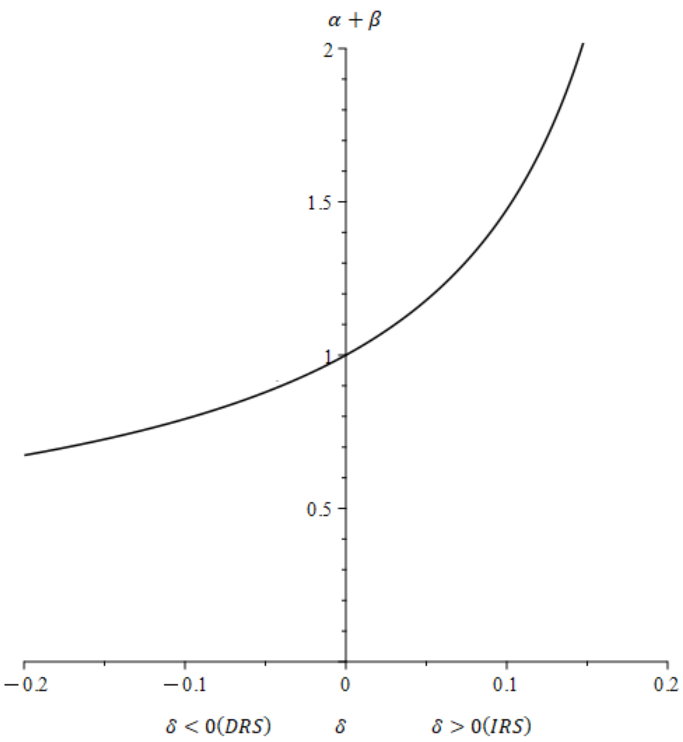


Figure A1: One-to-one correspondence between parameters δ and $\alpha+\beta$

The Producer Surplus for country i, denoted as π_i, and its partial derivative with respect to t_i are given as:

$$\pi_i(f) = p_i(f)x_{ii}(f) + \sum_{p\in M, z\neq i} p_p(f)x_{ip}(f) + p_k(f)x_{ik}(f) + \sum_{z\in N, z\neq k} p_z(f)x_{iz}(f) - c\Bigg(x_{ii}(f) + \sum_{p\in M, z\neq i} x_{ip}(f) + x_{ik}(f) + \sum_{z\in N, z\neq k} x_{iz}(f)\Bigg) + \delta\left(x_{ii}(f) + \sum_{p\in M, z\neq i} x_{ip}(f) + x_{ik}(f) + \sum_{z\in N, z\neq k} x_{iz}(f)\right)^2 - \sum_{p\in M, z\neq i} \tau x_{ip}(f) - t_k^f x_{ik}(f) - \sum_{z\in N, z\neq k} t_z^f x_{iz}(f)$$

$$\frac{d\pi_i(f)}{dt_i} = \frac{dp_i(f)}{dt_i}x_{ii}(f) + p_i(f)\frac{dx_{ii}(f)}{dt_i} + \sum_{p\in M, p\neq i}\left(\frac{dp_p(f)}{dt_i}x_{ip}(f) + \Big(p_p(f) - \tau\Big)\frac{dx_{ip}(f)}{dt_i}\right) + \frac{dp_k(f)}{dt_i}x_{ik}(f) + \Big(p_k(f) - t_k^f\Big)\frac{dx_{ik}(f)}{dt_i} + \sum_{z\in N, z\neq k}\left(\frac{dp_z(f)}{dt_i}x_{iz}(f) + (p_z(f) - t_z^f)\frac{dx_{iz}(f)}{dt_i}\right) + \left(\sum_{p\in M}\frac{dx_{ip}(f)}{dt_i} + \sum_{z\in N}\frac{dx_{iz}(f)}{dt_i}\right)\left(-c + 2\delta\Big(\sum_{p\in M,} x_{ip}(f) + \sum_{z\in N} x_{iz}(f)\Big)\right)$$

Now all member countries apply t_m^f, and non-member countries apply t_n^f for their optimum external tariffs. We obtain the x_s at equilibrium.

$$x_{ks}^*(f) = \eta + \frac{(\xi_{ks}(t_k^f) - (n-m-2)\xi(t_N) - \xi_{ks}(t_s^f)) \cdot t_n - \xi(t_{NM})t_m \cdot m + \tau\xi_{ks}(\tau)}{\xi}$$

$$x_{ki}^*(f) = \eta + \frac{(\xi_{ki}(t_k^f) - (n-m-1)\xi(t_N))t_n - ((m-1)\xi(t_{NM}) + \xi_{ki}(t_i^f))t_m + \tau\xi_{ki}(\tau)}{\xi}$$

$$x_{kk}^*(f) := \eta + \frac{(\xi_{kk}(t_k^f) - (n-m-1)\xi(t_N))t_n - \xi(t_{NM})t_m m + \tau\xi_{kk}(\tau)}{\xi}$$

$$x_{ik}^*(f) = \eta + \frac{(-(n-m-1)\xi(t_N) - \xi_{ik}(t_k^f))t_n + \xi(t_{MM})t_m m - \tau\xi(\tau)}{\xi}$$

$$x_{ii}^*(f) = \eta + \frac{((m-1)\xi(t_{MM}) + \xi_{ii}(t_i^f))t_m - t_n(n-m)\xi(t_N) - \tau\xi_{ii}(\tau)}{\xi}$$

$$x_{ip}^*(f) = \eta + \frac{((m-1)\xi(t_{MM}) + \xi_{ij}(t_j^f))t_m - t_n(n-m)\xi(t_N) - \tau\xi_{ij}(\tau)}{\xi}$$

To obtain the best response function for member country i's welfare with respect to the non-number of countries k, we sum the first-order derivatives for consumer surplus CS_i, profit π_i, and tariff revenue function TR_i:

$$\frac{dW_i}{dt_m^f}(f) = \frac{CS_i}{dt_m^f}(f) + \frac{d\pi_i}{dt_m^f}(f) + \frac{dTR_i}{dt_m^f}(f)$$
$$\implies \frac{dW_i}{dt_m^f}(f) = \psi_1 t_m^f + \psi_2 t_n^f + \psi_3$$

Likewise, to obtain the best response function for nonmember country k's welfare with respect to the number of countries i, we sum the first-order derivatives for consumer surplus CS_k, profit π_k, and tariff revenue function TR_k:

$$\frac{dW_k}{dt_n^f}(f) = \frac{CS_k}{dt_n^f}(f) + \frac{d\pi_k}{dt_n^f}(f) + \frac{dTR_k}{dt_n^f}(f)$$

$$\implies \frac{dW_k}{dt_m^f}(f) = \psi_4 t_m^f + \psi_5 t_n^f + \psi_6$$

where

$$\psi_1 = \frac{1}{\xi^2}\Bigg[(2\delta-1)\xi_{ii}(t_i^f)^2 + 2\Big(((n-1)(m+1)\delta - m+1)\xi(t_{MM}) + (m-1)\xi_{ij}(t_j^f)\delta\Big)\xi_{ii}(t_i^f) + (m-1)\Big($$

$$(m-1)\xi_{ij}(t_j^f) + ((m-2)^2 + 2(n-1)\delta - 2)\xi(t_{MM}) + (m-n)\left((m-2)\xi(t_{NM}) + 2\xi_{ki}(t_i^f)\right)\Bigg)\xi_{ij}(t_j^f)$$

$$+ (m-n)\Big((m-n)\xi_{ki}(t_i^f) + (m-1)\big((m-2)\xi(t_{MM}) + (m-n)\xi(t_{NM})\big) + 2\xi\Big)\xi_{ki}(t_i^f) - 2\Big($$

$$(-2+n)(-\delta n+m) - \delta + 1\Big)m\xi(t_{MM})^2 - (m-n)\Big(2(1+(-2+n)m)\xi(t_{MM}) - (m-1)\xi\Big)\xi(t_{NM})\Bigg]$$

$$\psi_2 = \frac{(m-n)}{\xi^2}\Bigg[\Big(-(1+2(-n+1)\delta)\xi_{ii}(t_i^f) - (2(n-1)((-n+1)\delta+m) - 2m)\xi(t_{MM})$$

$$+ (m-n)((m-1)\xi_{ki}(t_i^f) - (-2+n)\xi(t_{NM})\Big)\xi(t_N) + (m-1)\Big(-\xi_{ki}(t_k^f) + (m-1)\xi(t_N)\Big)\xi_{ij}(t_j^f)$$

$$-\Big(((-2n+2)\delta + 2m)\xi(t_{MM}) + (m-n)\xi(t_{NM}) - 2\delta\xi_{ii}(t_i^f)\Big)\xi_{ik}(t_k^f) - \Big((m-n)\xi_{ki}(t_i^f) - (m-1)\xi(t_{MM})$$

$$+\xi\Big)\xi_{ki}(t_k^f) + \xi(t_{MM})\Big(\xi + \xi_{kk}(t_k^f)\Big) + (n-m-1)\Big(\Big(\xi_{ks}(t_k^f) - \xi_{ks}(t_s^f)\Big)\xi(t_{MM}) + \Big((m-n)\xi_{ki}(t_i^f) + \xi$$

$$-(-2+n)\xi(t_{MM}) + (m-1)\xi_{ij}(t_j^f)\Big)\xi(t_N)\Bigg]$$

$$\psi_3 = \frac{1}{\xi^2}\Bigg[\Big((2\delta n-1)\xi_{ii}(t_i^f) + (m-1)(n-1)\xi_{ij}(t_j^f) + (m-n)\big((n-1)\left(\xi_{ki}(t_i^f) - \xi(t_{NM})\right) - \xi\big) - (n-1)\big($$

$$(-2\delta+1)n + m)\xi(t_{MM})\Bigg)\eta\xi + \tau\Bigg(-2(m-1)\delta\xi_{ij}(\tau) - (2\delta-1)\xi_{ii}(\tau) + 2\delta(m-n)\xi_{ik}(\tau)$$

$$+\frac{(\alpha-c)}{\tau}\xi\Big)\xi_{ii}(t_i^f) + (m-1)\tau\Bigg(\Big((-m+1)\xi_{ij}(t_j^f) + (n-m)\big(\xi_{ki}(t_i^f) - \xi(t_{NM})\big) + m\xi(t_{MM})\Big)\xi_{ij}(\tau)$$

$$+\big((n-m)\xi_{ki}(\tau) + \xi\big)\xi_{ij}(t_j^f)\Bigg) + (n-m)\tau\Bigg(\big((m-n)\xi_{ki}(t_i^f) - (m-1)\xi(t_{MM}) + \xi\big)\xi_{ki}(\tau) + (m-n)$$

$$\xi_{ik}(\tau)\xi(t_{NM})\Bigg) + \tau\Bigg(\big((-2n+2)\delta + m - 1\big)\big(\xi_{ii}(\tau) - (m-n)\xi_{ik}(\tau) + (m-1)\xi_{ij}(\tau)\big)$$

$$+\big((n-m-1)\xi_{ks}(\tau) + \xi_{kk}(\tau) - (m+1)\xi_{ik}(\tau)\big)(m-n)\Bigg) + \big((\alpha-c)(n-1) - (m-1)\tau\big)\xi\Bigg]$$

$$\psi_4 = \frac{m}{\xi^2}\Bigg[-2\Big((\delta m-1)\xi_{ki}(t_k^f)+\delta\xi_{kk}(t_k^f)\Big)\xi_{ki}(t_i^f) - \Big((m-1)\Big(m\xi(t_{MM})+(m-1)\xi(t_{NM})+\xi_{ij}(t_j^f)\Big)$$

$$+\Big(1+(2\delta-1)(n-1)m\Big)\xi(t_{NM})+(\xi+\xi_{ii}(t_i^f))\Big)\xi_{ki}(t_k^f)+\Big((m-n)-2\Big(-1+(n-1)\delta\Big)\xi(t_{NM})\Big)\xi_{kk}(t_k^f)$$

$$+(n-m-1)\Big((-m\xi(t_{MM})-2\delta\Big((n-1)\xi(t_{NM})+\xi_{ki}(t_i^f)\Big)\Big)\xi_{ks}(t_k^f)-\Big((m-n+1)\xi_{ks}(t_s^f)+\xi$$

$$+(m-n-1)\xi_{ks}(t_k^f)+\xi(t_N)(-2+n)-\xi_{kk}(t_k^f)+\xi_{ki}(t_k^f)\Big)\xi(t_{NM})-m\xi_{ks}(t_s^f)\xi(t_{MM})-\Big(\xi(t_N)-$$

$$\xi_{ki}(t_k^f)\Big)\xi_{ki}(t_i^f)\Bigg)-m\Bigg(\Big(\xi(t_{NM})(-2+n)+\xi_{ki}(t_i^f)\Big)\xi(t_N)-\Big(1-m\xi_{ik}(t_k^f)\Big)\xi(t_{MM})+$$

$$\Big((m-n)+1\Big)\xi(t_{NM})\xi_{ik}(t_k^f)\Bigg)\Bigg]$$

$$\psi_5 = \frac{1}{\xi^2}\Bigg[(n-m-1)\Bigg(m\Bigg(\Big(m\xi_{ik}(t_k^f)-(m-n)\xi_{ks}(t_k^f)-(m-n+2)\xi_{ks}(t_s^f)+\xi(t_N)(-n+2)-\xi\Big)\xi(t_N)$$

$$-\Big(-2\xi_{ks}(t_s^f)+(m-n+2)\xi(t_N)\Big)\xi_{ik}(t_k^f)+\Big(4\delta\xi_{ks}(t_k^f)-\Big(\Big(2n-2)\delta+m-n-2\Big)\xi(t_N)$$

$$+2\delta\xi_{ks}(t_s^f)\Big)\xi_{ki}(t_k^f)\Bigg)-\Bigg(2\Big(\delta(m-n+1)-\frac{m}{2}+\frac{n}{2}\Big)\xi_{ks}(t_k^f)-\Big(2(n-1)(m-n+1)\delta+(m-n)^2$$

$$-2\Big)\xi(t_N)+\Big(2\Big(\delta(m-n+1)-\frac{m}{2}+\frac{n}{2}\Big)\xi_{ks}(t_k^f)-\Big(2(n-1)(m-n+1)\delta+(m-n)^2-2\Big)\xi(t_N)$$

$$+2\Big(\frac{1}{2}+\delta(m-n+1)\Big)\xi_{ks}(t_s^f)+(1-4\delta)\xi_{kk}(t_k^f)\Bigg)\xi_{ks}(t_k^f)+\Bigg((n-m-1)\Big(\xi_{ks}(t_s^f)+\Big(n-m-3\Big)\xi(t_N)\Big)$$

$$-2\Big(\delta\xi_{kk}(t_k^f)+\xi\Big)\Bigg)\xi_{ks}(t_s^f)+\xi(t_N)\Big((m-n+2)\xi+\Big(1-2(n-1)\delta\Big)\xi_{kk}(t_k^f)+(-2+n)(m-n+1)\xi(t_N)$$

$$+m\Bigg(\xi_{ik}(t_k^f)\Big(m\xi_{ik}(t_k^f)-2\xi\Big)-\xi_{ki}(t_k^f)\Big(\Big(m\xi(t_N)-\xi_{ki}(t_k^f)\Big)(m-n-1)-2\delta(m\xi_{ki}(t_k^f)+2\xi_{kk}(t_k^f))\Big)\Bigg)$$

$$+\xi_{kk}(t_k^f)^2(2\delta-1)\Bigg]$$

$$\psi_6 = \frac{1}{\xi^2}\Bigg[(n-m-1)\Bigg(\tau\Bigg(\Big((n-m)(2\delta-1)-2\delta\Big)\xi_{ks}(t_k^f) - (n-m-1)\Big(\xi_{ks}(t_s^f) - \xi(t_N)\Big) + 2\delta\xi_{kk}(t_k^f) + \xi$$

$$+ m\Big(\xi(t_N) - \xi_{ik}(t_k^f) + 2\delta\xi_{ki}(t_k^f)\Big)\Bigg)\xi_{ks}(\tau) + \Bigg(\Big(-1+(2\delta-1)n\Big)\eta\xi + \tau\Big(m\Big(\xi_{ik}(\tau) + 2\delta\xi_{ki}(\tau)\Big)$$

$$+ (2\delta-1)\xi_{kk}(\tau)\Big) + (\alpha-c)\xi\Bigg)\xi_{ks}(t_k^f) - \Big((n-1)\eta\xi - m\tau\xi_{ik}(\tau)\Big)\xi_{ks}(t_s^f) + \Big((n-1)\eta\xi + m\tau\xi_{ki}(\tau)\Big)\xi(t_N)\Bigg)$$

$$+ \Big((n-1)\Big(m\Big(\xi(t_N) - \xi_{ik}(t_k^f)\Big) + \xi\Big) + (2\delta n-1)\Big(m\xi_{ki}(t_k^f) + \xi_{kk}(t_k^f)\Big) - mn\xi_{ki}(t_k^f)\Big)\eta\xi$$

$$+ m\tau\Bigg(\Big((m(2\delta+1)-n-1)\xi_{ki}(t_k^f) + 2\delta\xi_{kk}(t_k^f) + m\xi(t_N)\Big)\xi_{ki}(\tau) + \Big(2\delta\xi_{kk}(\tau) + \xi_{ii}(\tau)$$

$$+ (m-1)\xi_{ij}(\tau)\Big)\xi_{ki}(t_k^f) + (m\xi_{ik}(t_k^f) - \xi)\xi_{ik}(\tau)\Bigg) + 2\Big(\delta - \frac{1}{2}\Big)\tau\xi_{kk}(t_k^f)\xi_{kk}(\tau) + (\alpha-c)\Big(\xi_{kk}(t_k^f) + m\xi_{ki}(t_k^f)\Big)$$

and

$\xi_{ks}(t_k^f) = 2\delta(n-1)(2+(-2\delta+1)n)$, $\xi_{ks}(\tau) = 2m(m-1)\delta(2+n(1-2\delta))$,
$\xi_{ks}(t_s^f) = 2(-2\delta n+1)(n+1-2(n-1)\delta)+2\delta(n+1)$, $\xi_{ii}(t_i^f) = (n-m)(2\delta(n-1)((2\delta-1)n-2)+n+1)$,
$\xi_{ii}(\tau) = (m-1)((2n-m+1)(2\delta n-1)(1-2\delta)+(n-m)(1+2\delta))$,
$\xi_{ij}(t_j^f) = (n-m)(2\delta(-2\delta n-n+2)+2n^2\delta(2\delta-1)+n+1)$
$\xi_{ij}(\tau) = 4n\delta^2((n-m+1)^2+n) - 2\delta((2n+1)(n+1)+(n-m)((n-m+1)(n+2)+1)) + (n-m+2)(n+1)$, $\xi_{ik}(\tau) = 2(m-1)\delta((n+1)^2 - 2n\delta(n-m+1) - m(n+2))$,
$\xi_{ik}(t_k^f) = 2(-2\delta n+1)(n+1-2(n-1)\delta)+2\delta(n+1)$, $\xi_{kk}(\tau) = 2m(m-1)\delta(2+(-2\delta+1)n)$,
$\xi_{kk}(t_k^f) = (n-1)(2\delta(-2\delta n-n+2)+2n^2\delta(2\delta-1)+n+1)$,
$\xi_{ki}(\tau) = (n+1)(m-1) - 2(n-m)(m-1)\delta(2+(-2\delta+1)n)$, $\xi_{ki}(t_k^f) = 2\delta(n-1)(2+(-2\delta+1)n)$,
$\xi = (n+1)(-2\delta n+1)(1+(-2\delta+1)n)$, $\xi_{t_{MM}} = 2\delta(n-m)(2-(2\delta-1)n)$, $\eta = \frac{\alpha-c}{1+(-2\delta+1)n}$,
$\xi_{t_N} = 2\delta(n(1-4\delta)+3)$, $\xi_{t_{NM}} = 2\delta(m(n+1)+(-2\delta n+1)(m+1))$,
$\xi_{ki}(t_i^f) := (n-m)(2\delta(n(2\delta+1)-2)+2n^2\delta(2\delta-1)-n-1)+(n+1)(2(n-1)\delta-1)(-1+(2\delta-1)n)$

The best responses are found as follows:

$$\frac{dW_i}{dt_m^f} = 0 \quad \Rightarrow \quad \psi_1 t_m^f + \psi_2 t_n^f + \psi_3 = 0$$

and

$$\frac{dW_k}{dt_n^f} = 0 \quad \Rightarrow \quad \psi_4 t_m^f + \psi_5 t_n^f + \psi_6 = 0$$

Solving both the best response functions for members and non-members, we obtai the following tariffs:

$$t_m^f = \frac{\psi_2\psi_6 - \psi_3\psi_5}{\psi_1\psi_5 - \psi_2\psi_4}, \quad t_n^f = -\frac{\psi_1\psi_6 - \psi_3\psi_4}{\psi_1\psi_5 - \psi_2\psi_4}$$

9.3 Optimum Tariffs under No Agreement:

Under no agreement, all countries impose their optimum external MFN tariffs t_i^ϕ. First, we obtain the following expressions that would be useful in deriving optimum tariffs.

$$x_{ii}^*(\phi) = \eta + \frac{\Big(\xi_{ii}(t_i^\phi)) - (n-1)\xi(t_N)\Big)t_i^\phi}{\xi}$$

$$x_{iz}^*(\phi) = \eta + \frac{\Big(\xi_{iz}(t_i^\phi) - \xi_{ij}(t_j^\phi) - (n-2)\xi(t_N)\Big)t_i^\phi}{\xi}$$

$$\xi_{ii}(t_i^\phi) = 2(n-1)^2\delta(n(2\delta-1)-2) + (n-1)(n+1)$$

$$\xi_{iz}(t_i^\phi) = 2(n-1)\delta(n+2(1-n\delta))$$

$$\xi_{ij}(t_j^\phi) = [2(1-2\delta n)(n+1-2\delta(n-1)) + 2\delta(n+1)]$$

In order to find optimum external MFN tariffs under no agreement:

$$\frac{dW_i}{dt_i^\phi}(\phi) = \frac{CS_i}{dt_i^\phi}(\phi) + \frac{d\pi_i}{dt_i^\phi}(\phi) + \frac{dTR_i}{dt_i^\phi}(\phi)$$

$$\implies \frac{dW_i}{dt_i^\phi}(\phi) = \psi_7 t_i^\phi + \psi_8$$

As a result, we obtain

$$t_i^\phi = \frac{-\psi_8}{\psi_7}$$

where

$$\psi_7 = \frac{1}{\xi^2}\Bigg[(n-1)\left(\left(\xi(t_N)\left(n(n-1)-1\right) - \left(\xi_{iz}(t_i^\phi) - \xi_{ii}(t_i^\phi)\right)\right)\xi_{iz}(t_i^\phi)\right. - (2\delta(n-1)-1)\Bigg(\Bigg(\xi_{iz}(t_i^\phi)(n-1)$$
$$+ \xi_{ii}(t_i^\phi)\Bigg)\xi(t_N) + \xi_{iz}(t_i^\phi)\xi_{ij}(t_j^\phi)\Bigg) - \left((n-1)\left(\xi(t_N) - \xi_{ij}(t_j^\phi)\right) + \xi\right)\left(\xi(t_N)(n-2) + \xi_{ij}(t_j^\phi)\right)$$
$$- \xi_{ij}(t_j^\phi)\left(2\xi_{ii}(t_i^\phi)\delta + \xi\right)\Bigg) + \left(\xi_{iz}(t_i^\phi)(n-1) + \xi_{ii}(t_i^\phi)\right)^2(-1+2\delta)\Bigg]$$

$$\psi_8 = \frac{1}{\xi}\Bigg[(\eta+(n-1)\eta)\Big((n-1)\Big((-1+2\delta)\xi_{iz}(t_i^{\phi})-\xi_{ij}(t_j^{\phi})\Big)+2\delta\xi_{ii}(t_i^{\phi})\Big)+\eta\Bigg((n-1)\Big(\xi(t_N)(n-1)+\xi$$
$$-\xi_{iz}(t_i^{\phi})+\xi_{ij}(t_j^{\phi})\Big)-\xi_{ii}(t_i^{\phi})\Bigg)+(\alpha-c)\Big(\xi_{iz}(t_i^{\phi})(n-1)+\xi_{ii}(t_i^{\phi})\Big)\Bigg]$$

9.4 Pre-FTA MFN Tariff Reduction Trends:

One potential concern with our results is the possibility of pre-existing MFN tariff reduction trends in the years prior to FTA formation. In particular, if industries characterized by IRS were already experiencing faster declines in MFN tariffs before the FTA, then the post-FTA tariff cuts we observe might simply reflect a continuation of those trends rather than a causal effect of the FTA. To address this concern, we conduct a placebo test using the same regression framework as in Equation (8) of the main text, but we replace the dependent variable with the change in MFN tariffs before the FTA. Specifically, for each FTA, we calculate the MFN tariff change in the K years leading up to the FTA (for $K = 1$, 2, and 3) and regress it on the same set of regressors as in Equation (8): the cumulative FTA tariff preference (the negotiated preferential tariff reduction T years after FTA formation, interacted with the FTA partners' import share) along with the same set of fixed effects. We estimate these regressions for horizons $T = 1$ through 6, mirroring the main analysis, to examine wheather future FTA tariff preferences are correlated with prior MFN tariff changes.

The results in Table A1 show no evidence of differential pre-FTA trends. Column (1) presents estimates from the benchmark specification, which does not distinguish industries by scale economies. The estimated coefficients on the FTA preference variable are uniformly negative across all horizons $T = 1, ..., 6$ and largely insignificant. This pattern holds for each pre-FTA window ($K = 1$, 2, or 3 years prior). Accordingly, there is no indication that products which eventually received larger FTA tariff reductions were on a trajectory of faster MFN tariff cuts in advance of the agreement. This null finding on the pre-trend test implies that future FTA tariff concessions do not appear to be "predicted" by prior MFN liberalization.

Furthermore, we find no systematic difference between IRS and DRS industries in their pre-FTA tariff trajectories. The regression allowing for IRS-DRS differences shows that the estimated coefficient gap between IRS and DRS sectors (see the final column of Table A1 is almost never statistically significant. This indicates that IRS industries were not on a markedly

Table A1: Estimation results for model (E8): Pre-trends / placebo (MFN tariff changes before FTA)

	(1) Baseline	(2) β_T	(3) β_T^I	(4) β_T^D	(5) $\beta_T + \beta_T^I$	(6) $\beta_T + \beta_T^D$	(7) $\beta_T^I - \beta_T^D$
				MFN Tariff Change in 1 Year Before FTA			
PM(T=1)×PS	-1.535***	-1.731***	-0.910	1.382***	-2.641**	-0.349**	-2.292**
	(0.385)	(0.363)	(1.097)	(0.349)	(1.100)	(0.167)	(1.117)
PM(T=2)×PS	-0.528***	-0.795***	0.579	0.820***	-0.216	0.025	-0.241
	(0.087)	(0.112)	(0.540)	(0.161)	(0.505)	(0.089)	(0.455)
PM(T=3)×PS	-0.181	-0.200	-0.035	0.066	-0.235	-0.134	-0.101
	(0.175)	(0.220)	(0.363)	(0.192)	(0.465)	(0.096)	(0.428)
PM(T=4)×PS	-0.114	-0.124	-0.660	0.054	-0.784	-0.070	-0.714
	(0.136)	(0.216)	(0.633)	(0.171)	(0.818)	(0.065)	(0.772)
PM(T=5)×PS	-0.379	-0.378	-0.559	0.215	-0.937	-0.163	-0.774
	(0.300)	(0.306)	(0.694)	(0.172)	(0.758)	(0.254)	(0.618)
PM(T=6)×PS	-0.262	-0.397	0.200	0.346	-0.197	-0.051	-0.146
	(0.229)	(0.266)	(0.189)	(0.213)	(0.202)	(0.059)	(0.177)
				MFN Tariff Change in 2 Years Before FTA			
PM(T=1)×PS	-1.798***	-2.053***	0.575	1.200*	-1.479*	-0.853	-0.625
	(0.456)	(0.370)	(0.809)	(0.642)	(0.822)	(0.849)	(1.323)
PM(T=2)×PS	-0.476***	-0.779***	0.863***	0.848***	-1.479*	-0.853	-0.625
	(0.076)	(0.088)	(0.171)	(0.146)	(0.822)	(0.849)	(1.323)
PM(T=3)×PS	-0.005	-0.050	0.116	0.108	0.066	0.058	0.008
	(0.277)	(0.360)	(0.328)	(0.274)	(0.476)	(0.132)	(0.382)
PM(T=4)×PS	0.025	0.064	-0.356	-0.037	-0.293	0.026	-0.319
	(0.152)	(0.288)	(0.488)	(0.231)	(0.689)	(0.082)	(0.658)
PM(T=5)×PS	-0.557***	-0.578***	-0.255	0.388***	-0.833	-0.190	-0.643
	(0.147)	(0.121)	(0.593)	(0.147)	(0.676)	(0.248)	(0.509)
PM(T=6)×PS	-0.312	-0.489***	0.305***	0.448***	-0.184	-0.041***	-0.143
	(0.191)	(0.122)	(0.037)	(0.111)	(0.128)	(0.014)	(0.117)
				MFN Tariff Change in 3 Years Before FTA			
PM(T=1)×PS	-1.727***	-1.872***	-0.433	0.934	-2.305*	-0.938	-1.367
	(0.442)	(0.377)	(1.253)	(0.718)	(1.280)	(0.836)	(1.581)
PM(T=2)×PS	-0.446***	-0.720***	0.545	0.832***	-0.175	0.112	-0.287
	(0.103)	(0.098)	(0.507)	(0.178)	(0.503)	(0.170)	(0.434)
PM(T=3)×PS	-0.075	-0.128	0.011	0.167	-0.117	0.039	-0.156
	(0.266)	(0.347)	(0.363)	(0.298)	(0.476)	(0.132)	(0.423)
PM(T=4)×PS	-0.062	-0.093	-0.613	0.082	-0.706	-0.010	-0.696
	(0.133)	(0.246)	(0.597)	(0.219)	(0.787)	(0.072)	(0.772)
PM(T=5)×PS	-0.334	-0.347	-0.472	0.375*	-0.818	0.028	-0.846
	(0.332)	(0.341)	(0.689)	(0.212)	(0.742)	(0.265)	(0.627)
PM(T=6)×PS	-0.284	-0.450	0.091	0.443*	-0.359	-0.007	-0.352
	(0.261)	(0.286)	(0.187)	(0.249)	(0.254)	(0.075)	(0.236)

Notes: The dependent variable is the cumulative change in the MFN tariff over one year before the FTA (top panel), two years before the FTA (middle panel), and three years before the FTA (bottom panel). * significant at 10%, ** significant at 5%, *** significant at 1%. Standard errors, shown in parentheses, are calculated at the country levels. All specifications include country and HS6 product fixed effects.

different MFN liberalization path than DRS industries prior to FTA implementation. Taken together, across all pre-FTA windows and horizons, the results provide no evidence that future FTA tariff preferences were anticipated by, or correlated with, earlier MFN tariff cuts. These placebo tests therefore reinforce the interpretation that the post-FTA tariff complementarity effects documented in the main text are not driven by pre-existing trends.

References

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

[2] Bagwell, K., and Staiger, R. (1997a) "Multilateral Tariff Cooperation during the Formation of Customs Unions." *Journal of International Economics*, 42(1), 91–123.

[3] Bagwell, K., and Staiger, R. (1997b) "Multilateral Tariff Cooperation during the Formation of Free Trade Areas." *International Economic Review*, 38(2), 291–319.

[4] Bagwell, K., and Staiger, R. (1998) "Regionalism and Multilateral Tariff Cooperation." *International Trade Policy and the Pacific Rim.* London: Macmillan.

[5] Bagwell, K., and Staiger, R. (2005a) "Multilateral Trade Negotiations, Bilateral Opportunism, and the Rules of GATT." *Journal of International Economics*, 67(2), 268–294.

[6] Bagwell, K., and Staiger, R. (2005b) "Enforcement, Private Political Pressure, and the GATT/WTO Escape Clause." *Journal of Legal Studies*, 34(2), 471–513.

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

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

[9] Bi, Z., and Stoyanov, A. (2024) "Integrated Analysis of the Effect of Free Trade Agreements on Multilateral Trade Liberalization." SSRN Working Paper No. 5268488.

[10] Bond, E. W., Riezman, R., and Syropoulos, C. (2004) "A Strategic and Welfare Theoretic Analysis of Free Trade Areas." *Journal of International Economics*, 64, 1–27.

[11] Bond, E. W. (2009) "Paths of Efficient Self-Enforcing Trade Agreements." *Economic Theory*, 41(1), 85–104.

[12] Brander, J., and Krugman, P. (1983) "A 'Reciprocal Dumping' Model of International Trade." *Journal of International Economics*, 15(3), 313–321.

[13] Calvo-Pardo, H., Freund, C., and Ornelas, E. (2009) "The ASEAN Free Trade Agreement: Impact on Trade Flows and External Trade Barriers." *CEP Discussion Papers*, dp0930, Centre for Economic Performance, LSE.

[14] Chi, E. C., and Yildiz, H. M. (2024a) "Bilateralism vs. Multilateralism: Role of Political Economy for the Prospect of Multilateral Free Trade." *Canadian Journal of Economics/Revue canadienne d'économique*, 57(3), 870–900.

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

[16] Estevadeordal, A., Freund, C., and Ornelas, E. (2008) "Does Regionalism Affect Trade Liberalization Towards Nonmembers?" *The Quarterly Journal of Economics*, 123(4), 1531–1675.

[17] Freund, C., and McDaniel, C. (2016) "How Long Does it Take to Conclude a Trade Agreement with the US?" *Trade & Investment Policy Watch*. Peterson Institute for International Economics.

[18] Furusawa, T., and Konishi, H. (2007) "Free Trade Networks." *Journal of International Economics*, 72(2), 310–335.

[19] Goyal, S., and Joshi, S. (2006) "Bilateralism and Free Trade." *International Economic Review*, 47(3), 749–778.

[20] Krishna, P. (1998) "Regionalism and Multilateralism: A Political Economy Approach." *Quarterly Journal of Economics*, 113, 227–251.

[21] Krugman, P. R. (1991) "The Move Toward Free Trade Zones." In *Policy Implications of Trade and Currency Zones: A Symposium Sponsored by the Federal Reserve Bank of Kansas City*, pp. 7–41. Federal Reserve Bank of Kansas City, Kansas City.

[22] Kuenzel, D. J., and Sharma, R. R. (2021) "Preferential Trade Agreements and MFN Tariffs: Global Evidence." *European Economic Review*, 138, 103850.

[23] Lake, J. (2017) "Free Trade Agreements as Dynamic Farsighted Networks." *Economic Inquiry*, 55(1), 31–50.

[24] Lake, J., and Yildiz, H. M. (2016) "On the Different Geographic Characteristics of Free Trade Agreements and Customs Unions." *Journal of International Economics*, 103, 233–233.

[25] Lake, J., and Krishna, P. (2019) "Preferential Trade Agreements: Recent Theoretical and Empirical Developments." *Oxford Research Encyclopedia of Economics and Finance.*

[26] Lake, J., Nken, M., and Yildiz, H. M. (2020) "Tariff bindings and the dynamic formation of Preferential Trade Agreements." *Journal of International Economics*, 103279.

[27] Lashkaripour, A., and Lugovskyy, V. (2023) "Profits, Scale Economies, and the Gains from Trade and Industrial Policy." *American Economic Review*, 113(10), 2759–2808.

[28] Limão, N. (2006) "Preferential Trade Agreements as Stumbling Blocks for Multilateral Trade Liberalization: Evidence for the US." *American Economic Review*, 96(3), 896–914.

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

[30] Mai, J., and Stoyanov, A. (2015) "The Effect of the Canada–US Trade Agreement on Canadian Multilateral Trade Liberalization." *Canadian Journal of Economics*, 48(3), 1067–1098.

[31] McLaren, J. (2002) "A Theory of Insidious Regionalism." *Quarterly Journal of Economics*, 117(2), 571–608.

[32] Missios, P., Saggi, K., and Yildiz, H. M. (2016) "External Trade Diversion, Exclusion Incentives and the Nature of Preferential Trade Agreements." *Journal of International Economics*, 99, 105–119.

[33] Mölders, F. (2012) "On the Path to Trade Liberalization: Political Regimes in International Trade Negotiations." Mimeo.

[34] Mölders, F. (2015) "On the Path to Trade Liberalization: Political Regimes in Trade Negotiations." *World Economy*, 39(7), 890–924.

[35] Nken, M., and Yildiz, H. M. (2022) "Implications of Multilateral Tariff Bindings on the Extent of Formation of Preferential Trade Agreements." *Economic Theory*, 73(1), 301–347.

[36] Odell, J. (2006) *Negotiating Trade: Developing Countries in the WTO and NAFTA.* Cambridge University Press.

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

[38] Ornelas, E. (2007) “Exchanging Market Access at the Nonmember’s Expense: The Case of Customs Unions.” *Canadian Journal of Economics*, 40(1), 207–224.

[39] Ornelas, E. (2008) “Feasible Multilateralism and the Effects of Regionalism.” *Journal of International Economics*, 74(1), 202–224.

[40] Ornelas, E., and Liu, X. (2012) “Free Trade Agreements and the Consolidation of Democracy.” Mimeo.

[41] Riezman, R. (1999) “Can Bilateral Trade Agreements Help Induce Free Trade?” *Canadian Journal of Economics*, 32, 751–766.

[42] Saggi, K. (2006) “Preferential Trade Agreements and Multilateral Tariff Cooperation.” *International Economic Review*, 47, 29–57.

[43] Saggi, K., and Yildiz, H. (2010) “Bilateralism, Multilateralism, and the Quest for Global Free Trade.” *Journal of International Economics*, 81(1), 26–37.

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

[45] Saggi, K., Woodland, A., and Yildiz, H. M. (2013) “On the Relationship Between Preferential and Multilateral Trade Liberalization: The Case of Customs Unions.” *American Economic Journal: Microeconomics*, 5(1), 63–99.

[46] Saggi, K., Stoyanov, A., and Yildiz, H. M. (2018) “Do Free Trade Agreements Affect Tariffs of Non-member Countries? A Theoretical and Empirical Investigation.” *American Economic Journal: Applied Economics*, 10(3), 128–170.

[47] Saggi, K., Wong, W. F., and Yildiz, H. M. (2019) “Should the WTO Require Free Trade Agreements to Eliminate Internal Tariffs?” *Journal of International Economics*, 118, 316–330.

[48] Seidmann, D. (2009) "Preferential Trading Arrangements as Strategic Positioning." *Journal of International Economics*, 79, 143–159.

[49] Stoyanov, A., and Yildiz, H. M. (2015) "Preferential versus Multilateral Trade Liberalization and the Role of Political Economy." *European Economic Review*, 80, 140–164.

[50] Xue, L., Zhang, J., and Zu, L. (2013) "Farsighted Free Trade Networks." *International Journal of Game Theory*, 42(2), 375–398.

[51] Yi, S. (1996) "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good." *Journal of International Economics*, 41(1), 153–177.