Snowballing alongside Domino on Proliferation of Preferential Trade Agreements

Jee-Hyeong Park
Seoul National University

Jaeyoun Roh*
Yeungnam University

January 22, 2022

Abstract

With regard to the third-country effects on bilateral incentives to sign new preferential trade agreements (PTAs), we develop a theory-based empirical approach that enables the assessment of how a country's pre-existing PTAs affect her formation of a new PTA. Assessment of this effect is not trivial because a country's pre-existing PTAs are also the pre-existing PTAs of her potential PTA partner's partner, generating two counter-acting effects on the bilateral incentives to sign a new PTA. Our empirical analysis shows that a country's pre-existing PTAs generate a positive effect on her incentive to sign a new PTA (i.e., Own PTA effect) but they generate a negative effect on her new partner's incentive to sign a PTA (i.e., Partner's PTA effect), as our theory predicts. The analysis also reveals that members of a CU are less likely to sign a new PTA than non-CU member countries as the Partner's PTA effect is magnified by the CU's joint negotiation requirement for a new PTA. Based on our n-country model of PTA proliferation, we conduct some calibration exercises that yield a varying degree of predictive power.

*Park: Professor, Department of Economics, Seoul National University (j-hpark@snu.ac.kr). Roh: Assistant Professor, Department of International Trade, Yeungnam University (acts8@yu.ac.kr). For their useful comments, we would like to thank Pol Antras, Chul Chung, and the conference participants at the IEFS-EAER Conference at Seoul National University. We also thank Daeeun Bae and Taegon Yang for their research assistance. Roh acknowledges the financial support from the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019 1A 5A 8032141).

1 Introduction

The rapid proliferation of preferential trade agreements (PTAs) started since the early 1990s has drawn a lot of attention, generating numerous studies on various aspects of it. With regard to the potential cause and process of such proliferation, arguably the best-known economic analysis is "a domino theory of regionalism" by Baldwin (1993): The trade diversion effect of a PTA triggers a *third* (i.e., not-being a part of the PTA) country to pursue a new PTA to redress her negatively-affected domestic industry's profit, which in turn may trigger a PTA being pursued by another third country and so on. Baldwin and Jaimovich (2012) provide a supportive empirical evidence for this theory of PTAs being "contagious" based on a theory-based measure of *contagion*.

A related study that emphasizes *third-country* effects (i.e. effects of pre-existing PTAs with *third* countries that are not negotiating partners of new PTAs under consideration) on the formation of PTAs was conducted by Chen and Joshi (2010). According to their analysis based on a three-country trade model, a country's pre-existing PTA generates two opposite effects on the formation of a new PTA: A positive effect on her incentive to sign a PTA by reducing her new-PTA-associated welfare loss to be caused by decreases in her tariff revenue and her domestic firms' local profits from the increased import competition; and a negative effect on her trading partner's incentive to sign a PTA with her by reducing the partner's welfare gain to be expected from the exporting firms' profit increase from their improved market access.

The positive effect of a country's pre-existing PTA, namely the "loss-sharing effect" (in the sense that the exporting firms of her pre-existing PTA partner share the profit losses generated by a new PTA with her domestic firms in her local market), dominates the negative effect, namely the "concession-erosion effect" (in the sense that the expected increase in profits of exporting firms of a new PTA partner is eroded by competition from exporting firms of the pre-existing PTA partner) under the following three mutually-non-exclusive conditions: If the market size of a country with a pre-existing PTA relative to that of her new PTA partner is sufficiently large; if her production cost disadvantage relative to her new PTA partner's is sufficiently large; and if both PTA negotiating countries have pre-existing PTAs. The empirical analysis of Chen and Joshi (2010), employing "one-(of a pair of potential PTA partners has a pre-existing) PTA dummy" interacted with country-pair differences both in their GDPs and in their unit labor costs, as well as "two-(of a pair of potential PTA partners have pre-existing) PTA dummy", generates supportive evidence for these third-country effects of pre-existing PTAs on the formation of new PTAs.

Our paper contributes to the literature on PTA formation by analyzing the third-country effects of Chen and Joshi (2010) in the n-country trade model of Furusawa and Konishi (2007). This enables us to construct theory-based measures that distinguish the loss-sharing effect (which we call *Own PTA effect*, as they slightly differ from each other) from the concession-erosion effect (*Partner's PTA effect*) of a country's pre-existing PTAs, finding supportive empirical evidence for the concurrent existence of these two opposite forces on the formation of new PTAs. These theory-based measures of pre-existing PTA effects are closely related but also different from the *contagion* measure of Baldwin and Jaimovich (2012). Our regression analysis employing these measures of the pre-existing PTA effects (including the contagion measure) reveals that they all matters in formations of PTAs in a statistically significant way.

With regard to the proliferation of PTAs across the global economy, our analysis thus suggests that the *snowballing* effect of pre-existing PTAs alongside their domino effect have been playing an important role for it: Our Own PTA effect works like snowballing as it induces a country with a pre-existing PTA to sign an additional PTA, and such inducement repeats itself until no further PTA is signed by her. The Partner's PTA effect then plays the role of restraining this snowballing effect, putting some limit on a country's expansion of PTA partners by making a PTA less attractive for her potential new PTA partners.

Theoretically, note that these pre-existing PTA effects, some of them seemingly predicting opposite effects on the formation of new PTAs, may co-exist with each other in a more general model. The concession-erosion effect of a pre-existing PTA, supposedly generating a negative effect on signing a new PTA as suggested by Chen and Joshi (2010), may trigger a negatively-affected third country to pursue a new PTA with the country that caused such a negative effect according to Baldwin (1993) and Baldwin and Jaimovich (2012). This difference in their predictions on how a pre-existing PTA of a country affects her potential new PTA partner's incentive to sign a PTA with her largely stems from their modeling difference: The former focuses on the standard welfare consideration on a country's PTA formation decision, ignoring the political economy aspect of it; and the latter largely focuses on the political economy aspect of PTA formation by putting zero weight on the standard welfare term in considering the effect of trade policy, thus in their construction of relevant measures for empirical analysis.[1] In a model that allows both the standard welfare and political-economy considerations to influence countries' PTA formation decision, it should be possible for

[1]Baldwin and Jaimovich (2012) do consider the case in which governments care only about

these pre-existing PTA effects to co-exist.[2]

Our use of an n-country trade model also enables us to calibrate the model to explain the formation of PTAs across the world. For the period between 1993 and 2017, we conduct calibrations of two extended models of ours, finding a set of parameter values of each model that perform the best in predicting the PTA formation in our data. To reflect the possibility of countries' having different preferences toward a PTA, first we introduce a country-specific critical value parameter such that a country would have an incentive to sign or keep a PTA if her PTA-driven welfare change is above this level. Then, we calibrate this extended model to find the set of country-specific critical value parameters that performs the best in predicting PTA formation. The best prediction success rate is 98% for no PTA being signed, and it is 59% for a PTA being signed or kept.

The second extension is to introduce country-pair specific critical value parameters into the model such that a pair of countries will form a PTA only when both countries' PTA-driven welfare changes are above their pair-specific critical values. For this calibration, we limit our data to pairs of countries whose PTA status changes during the period of 1993 to 2017, as the critical value parameters will automatically take sufficiently large numbers for pairs of countries having no PTA for the whole period, obtaining a perfect fit of the model for such cases. Under this second calibration, the best prediction success rate goes down to 90% for no PTA being signed, and it goes up to 92% for a PTA being signed or kept.

Because it is beyond the scope of this paper to provide a comprehensive discussion of the large literature on the formation of PTAs, we will only discuss works

the conventional welfare in the appendix of their paper, demonstrating that the contagion effect still exists as long as the initial most-favored-nation tariff level is not too high so that the tariff revenue loss from a PTA is not very important. Instead of the CES utility function employed by Baldwin and Jaimovich (2012) over differentiated products, we uses the quadratic utility model, another utility function used in the trade literature, such as Krugman et al. (1990) and Ottaviano et al. (2002). In addition, the mass of firms that produce differentiated products in each country is fixed in our model, which is different from the free entry model of Baldwin and Jaimovich (2012). These additional modeling differences also contribute to the difference in their predictions on how pre-existing PTAs affect PTA formation.

[2]One may consider introducing a political economy consideration into our model by having a government place a positive weight on the lobby for a PTA by domestic exporting firms (in addition to her standard welfare consideration) on her PTA formation decision. If the government places a stronger weight for the lobby triggered by her exporting firms to regain their market share in her potential PTA partner's market due to this partner's newly-formed PTA with another country than the lobby for the PTA without such a market share reduction, then all three pre-existing PTA effects may coexist in such a model. As a possible cause for such asymmetric weighting on the lobby for a PTA with and without an external market shrinkage shock, one may consider a conservative social welfare function, such as the one suggested by Corden (1974).

that are closely related to our analysis.[3] With regard to the theoretical studies of the third-country effects of pre-existing PTAs on the proliferation of PTAs, Baldwin (1993), Baldwin and Jaimovich (2012), and Chen and Joshi (2010) are the most directly related works to ours, as discussed above. These studies and ours do consider the interdependence of bilateral PTAs on formation of PTAs, thus going beyond the basic approach of Baier and Bergstrand (2004) that treats each bilateral PTA as independent from each other. However, we also share these previous studies' limitation of focusing on each country's static incentive to sign a bilateral PTA with their pre-existing PTAs be simply given, ignoring countries' potential consideration of their optimal PTA formation path.

Despite this limitation, we do treat a country that belongs to a pre-existing customs union (CU) differently from the one that does not in analyzing her incentive to sign a bilateral PTA.[4] In analyzing the Partner's (pre-existing) PTA effect on a CU-member country's incentive to sign a PTA, we pretend that other members of her CU have a pre-existing PTA with her potential PTA partner. This is because a CU member country's improved access into her potential PTA partner's market from signing a PTA will be restrained by her CU's other members' having the same improved access due to the CU's common external trade policy requirement. Given the negative effect of a potential PTA partner's pre-existing PTAs on a country's incentive to sign a PTA with that partner, our model predicts that CU member countries are less likely to sign a PTA compared with non-CU member countries with the same characteristics. Our empirical analysis do supports this hypothesis, obtaining a larger negative coefficient estimate of the Partner's (pre-existing) PTA effect variable that is more robustly significant when we construct a CU member country's Partner's PTA effect variable as suggested above, comparing with the case of treating a CU member and a non-member in the same way for the construction of this variable.

This result on CU members' PTA formation incentive provides indirect empirical support for the following theoretical works on endogenous formation of PTAs that demonstrate the CU's less preferred aspect of unanimity-based membership expansion rule in attaining global free trade, compared with the bilateral-decision-based expansion of free trade agreements (FTAs). By analyzing the equilibrium

[3]See Limão (2016) for a recent survey of the literature on PTAs.

[4]Even though a CU has a common external trade policy so that it is supposed to negotiate a new PTA as one unit, we treat each member country of a CU as an independent decision maker (except that each has a pre-existing bilateral PTA with other CU members and shares the same external tariffs as her other CU members') with regard to her decision to sign a PTA in our empirical analysis, following the previous empirical studies on PTA formation.

structure of trading blocs, Yi (1996) demonstrates that the grand CU (i.e. global free trade) is an equilibrium outcome under the "open regionalism" rule but typically not under the "unanimous regionalism" rule.[5] Utilizing the network formation game of Jackson and Wolinsky (1996), Furusawa and Konishi (2007) show that the complete FTA network (attaining global free trade) is aunique stable network of symmetric countries if all PTAs take a form of FTA with a sufficiently low substitutability among differentiated goods or with sufficiently low external tariffs. In contrast, several CUs of asymmetric size may co-exist in a stable network if all PTAs take a form of CU.[6] More recently, Saggi et al. (2013), together with an analysis on FTAs by Saggi and Yildiz (2010), establish that the freedom to pursue PTAs may prevent the attainment of global free trade when such agreements take the form of CU but it may not prevent global free trade when they take the form of an FTA, by analyzing the Nash equilibria of a simultaneous-move trade agreement game between three countries.[7]

Our empirical analysis builds on the following studies on the formation of PTAs. We rely on Baier and Bergstrand (2004) in choosing the explanatory variables that do not involve pre-existing PTAs. To evaluate the impacts of pre-existing PTAs on the formation of PTAs, we largely follow the empirical specifications of Baldwin and Jaimovich (2012), which in turn follow Egger and Larch (2008) who extend the cross-sectional empirical choice model of Baier and Bergstrand (2004) into a panel setting and control for interdependence (between PTAs) by including a bilateral distance weighted measure of pre-existing PTAs. In addition to having our Own PTA effect and Partner's PTA effect variables, we also include the contagion measure of Baldwin and Jaimovich (2012) and this "interdependence" measure of Egger and Larch (2008) in our empirical analysis.[8]

This paper consists of following sections. Based on the n-country trade model of Furusawa and Konishi (2007), Section 2 analyzes net welfare changes caused

[5]The CU membership is "open" to all countries under the open regionalism rule, as long as they accept the CU's common external trade policy.

[6]Proposition 4 of Furusawa and Konishi (2007) also demonstrates that a pair of countries is less likely to have a free trade link if either of them is a member of a CU rather than a regional FTA for the symmetric-country case, with the condition for attaining the complete FTA network being satisfied.

[7]Even though these more recent studies also generate the result that global free trade is more likely to be attained under FTAs than under CUs, the driving force behind their result is not the unanimity requirement of a CU expansion. It is the tariff coordination of CU members that makes the incumbent members' exclusion (of a new member) incentive be stronger when the PTA takes a form of CU as opposed to a FTA, a mechanism that both Furusawa and Konishi (2007) and our analysis rule out by assuming that external tariffs of PTA members are given exogeneously.

[8]More recently, Baier et al. (2014) empirically show that "own-FTA" effects (owing to either

by signing or keeping a PTA, focusing on the pre-existing PTA effects. Section 3 provides an econometric analysis of our model's prediction on the formation of PTAs. In Section 4, we conduct calibrations of our model in explaining the formation of PTAs. Section 5 concludes.

2 Theoretical Model

2.1 Basic Set-up and Equilibrium Welfare

The basic set-up of our theoretical model follows Furusawa and Konishi (2007), having the world be consisted of n countries, populated by a continuum of identical consumers who consume a numeraire good and a continuum of differentiated industrial commodities. Each consumer has l units of labor, with each unit of labor producing one unit of the numeraire good under perfect competition, so that the wage rate is equal to 1. The unit labor requirement of the differentiated commodities is normalized to 0. Thus, all countries are identical in their capacity in producing differentiated commodities, which in turn implies that gains from trade come from expanding varieties.[9]

A differentiated commodity can be considered as a variety of an industrial good that is indexed by $\omega \in [0, 1]$. A differentiated industrial commodity, ω, is produced by one firm that belongs to one of n countries and engages in price competition in each country's market that is segmented from others'. In country $i \in N \equiv \{1, 2, ..., n\}$, there exist the measure μ^i of consumers and the measure s^i of firms (each produces one variety of an industrial good), both being normalized with $\sum_{k=1}^{n} \mu^k = 1$ and $\sum_{k=1}^{n} s^k = 1$. The ratio $\theta^i = s^i / \mu^i$, then measures country i's industrialization level. Finally, the mass of firms in each country, s^i, is exogeneously given, determining each country's potential market share in the global differentiated commodity market.[10]

PTA partner already having other FTAs) play a much bigger role than "cross-FTA" effects (owing to other FTAs existing in the rest of the world) on formation PTAs. While we do not analyze "cross-FTA" effects, our analysis shows that "snowballing" effects in addition to "domino" effects are the possible sources of their "own-FTA" effects.

[9]This simplifying assumption of Furusawa and Konishi (2007) enables us to assume way the potential loss from trade diversion in association with signing a PTA, focusing on the effects of pre-existing PTAs on formation of PTAs.

[10]Given the consumer preference specified below, if the price elasticity of a variety is equal to one, then s^i will be equal to the market share of country i in every country's differentiated commodity market.

Country i imposes a specific tariff at a rate of t_j^i on her import of differentiated commodities from country j. For simplicity, there is no commodity tax, having $t_i^i = 0$. In addition, we assume that each country does not impose any tariff on the numeraire good that are traded internationally to obtain trade balance. The tariff revenue is redistributed equally to domestic consumers.

Regarding pre-existing PTAs, let $C_i = \{k \in N \mid t_k^i = 0\}$ represent the set of countries that produce differentiated commodities on which country i imposes no tariffs, including country i herself. $\widehat{C}_i = \{k \in N_{-i} \mid t_k^i = 0\}$, thus excluding country i from C_i. Country i imposes a common Most Favored Nation (MFN) tariff on the non-PTA members' export of differentiated commodities, denoted by t^i. Then, country i's average tariff rate, denoted by $\bar{t}^i$, is equal to $(1 - s^{C_i})t^i$ with $s^{C_i} = \sum_{k \in C_i} s^k$.

A representative consumer's utility is given by the following quasi-linear utility function:

$$U(q, q_0) = \int_0^1 q(\omega)d\omega - \frac{1-\sigma}{2}\int_0^1 q(\omega)^2 d\omega - \frac{\sigma}{2}\left[\int_0^1 q(\omega)d\omega\right]^2 + q_0, \tag{1}$$

where $q(\omega)$ is the consumption of a variety $\omega \in [0, 1]$, q_0 is the consumption of the numeraire, and $\sigma \in (0, 1)$ represents the substitutability between the varieties. The consumer's utility maximization then yields the demand function for each variety as

$$q(\omega) = \frac{1}{1-\sigma}\left[1 - \widetilde{p}(\omega) - \sigma\left(1 - \widetilde{P}\right)\right] \tag{2}$$

where $\widetilde{P} = \int_0^1 \widetilde{p}(\omega)d\omega$, denotes the average consumer price for the industrial commodities in the consumer's country.

Firm ω of country k chooses $\{p_k^i(\omega)\}_{i=1}^n$ to maximize its profit

$$\pi(\omega) = \sum_{i=1}^n \mu^i p_k^i(\omega) q_k^i(\omega) \tag{3}$$

where $q_k^i(\omega) = \left[1 - p_k^i(\omega) - t_k^i - \sigma\left(1 - \widetilde{P}^i\right)\right] / (1-\sigma)$, with $p_k^i(\omega)$ and $\widetilde{P}^i$ denoting the producer price of a variety ω from country k and the average consumer price of the differentiated commodities in country i, respectively. The firm's profit maximization generates its producer price in country i as

$$p_k^i(\omega) = \frac{1}{2}\left[1 - t_k^i - \sigma\left(1 - \widetilde{P}^i\right)\right], \tag{4}$$

which in turn yields

$$\widetilde{P}^i = \frac{1-\sigma+\bar{t}^i}{2-\sigma} \tag{5}$$

from $\widetilde{P}^i = \sum_{k=1}^{n} s^k \left(p_k^i + t_k^i\right) = \frac{1}{2}\left[1+\bar{t}^i - \sigma\left(1-\widetilde{P}^i\right)\right]$.

The equilibrium producer price, p_k^i, that each firm of country k charges for country i's market, and the corresponding demand of a representative consumer, q_k^i, are then functions of country i's tariff vector, $\underline{t}^i = (t_1^i, ...t_n^i)$:

$$\begin{aligned} p_k^i(\underline{t}^i) &= \tfrac{1-\sigma}{2-\sigma} - \tfrac{1}{2}t_k^i + \tfrac{\sigma}{2(2-\sigma)}\bar{t}^i, \\ q_k^i(\underline{t}^i) &= \tfrac{1}{2-\sigma} - \tfrac{1}{2(1-\sigma)}t_k^i + \tfrac{\sigma}{2(1-\sigma)(2-\sigma)}\bar{t}^i. \end{aligned} \tag{6}$$

A representative consumer's income in country i is the sum of labor income, per capita tariff revenue, and per capita profit share of country i's firms:

$$\begin{aligned} y &= l + T^i(\underline{t}^i) + \tfrac{s^i \pi_i(\underline{t})}{\mu^i}, \text{ with} \\ T^i(\underline{t}^i) &= \sum_{k=1}^{n} t_k^i s^k q_k^i(\underline{t}^i), \text{ and} \\ \pi_i(\underline{t}) &= \sum_{k=1}^{n} \mu^k p_i^k(\underline{t}^k) q_i^k(\underline{t}^k) = \sum_{k=1}^{n} \mu^k (1-\sigma)\left[q_i^k(\underline{t}^k)\right]^2, \end{aligned} \tag{7}$$

where $\underline{t} = (\underline{t}^1, ...\underline{t}^n)$ denotes the world tariff vector. Then, from the budget constraint of $y = \int_0^1 \widetilde{p}(\omega)q(\omega)d\omega + q_0$, we can obtain

$$\begin{aligned} q_0 &= l + T^i(\underline{t}^i) + \tfrac{s^i\pi_i(\underline{t})}{\mu^i} - \sum_{k=1}^{n} s^k[p_k^i(\underline{t}^i) + t_k^i]q_k^i(t_k^i) \\ &= l + \sum_{k=1}^{n} t_k^i s^k q_k^i(\underline{t}^i) + \tfrac{s^i}{\mu^i}\sum_{k=1}^{n}\mu^k p_i^k(\underline{t}^k)q_i^k(\underline{t}^k) - \sum_{k=1}^{n} s^k[p_k^i(\underline{t}^i) + t_k^i]q_k^i(t_k^i) \\ &= l - \sum_{k\neq i} s^k p_k^i(\underline{t}^i)q_k^i(\underline{t}^i) + \tfrac{s^i}{\mu^i}\sum_{k\neq i}\mu^k p_i^k(\underline{t}^k)q_i^k(\underline{t}^k) = l - M^i\left(\underline{t}^i\right) + X^i\left(\underline{t}^{-i}\right), \end{aligned} \tag{8}$$

where $M^i\left(\underline{t}^i\right)$ and $X^i\left(\underline{t}^{-i}\right)$ respectively denote country i's per capita import and per capita export of differentiated commodities, with $\underline{t}^{-i}$ representing the world tariff vector excluding country i's tariff.

Now, solving for $q_k^i(\underline{t}^i)$ and substituting q_0 into the quasi-linear utility function, we can obtain a representative consumer's utility as a function of the world tariff vector as follows:

$$\begin{aligned} W^i(\underline{t}) &\equiv U(q_k^i(\underline{t}^i)_{k\in N}, q_0^i(\underline{t}^i)) = V^i(\underline{t}^i) + \left[X^i(\underline{t}^{-i}) - M^i(\underline{t}^i)\right], \\ \text{with } V^i(\underline{t}^i) &\equiv \sum_{k=1}^n s^k q_k^i(\underline{t}^i) - \tfrac{(1-\sigma)}{2}\sum_{k=1}^n \left[s^k q_k^i(\underline{t}^i)\right]^2 - \tfrac{\sigma}{2}\left[\sum_{k=1}^n s^k q_k^i(\underline{t}^i)\right]^2 + l, \end{aligned} \tag{9}$$

where $V^i(\underline{t}^i)$, namely *gross utility*, represents the utility from consuming the differentiated commodities plus l units of numeraire goods. As shown in (8), the consumption level of numeraire goods will be smaller (larger) than l by the trade deficit (surplus) in trading of differentiated commodities. The decomposition of welfare into this gross utility term (V) plus trade surplus in (9) is useful in analyzing the incentive to sign or keep a PTA.

2.2 Incentives to Sign a PTA

A PTA between countries i and j reduces or eliminates tariffs imposed on commodities imported from each other. PTA signing countries may change their MFN tariffs on non-PTA members, but we will assume that they keep their MFN tariffs at their original levels for simplicity. The welfare change in country i due to a PTA between countries i and j, then can be expressed as follows:

$$\begin{aligned} \triangle_j W^i &\equiv W^i(0, \underline{t}^i_{-j}; 0, \underline{t}^j_{-i}; \underline{t}^{-\{i,j\}}) - W^i(t^i_j, \underline{t}^i_{-j}; t^j_i, \underline{t}^j_{-i}; \underline{t}^{-\{i,j\}}) \\ &= \triangle_j V^i(\underline{t}^i) + \triangle_j \left[X^i(\underline{t}^{-i}) - M^i(\underline{t}^i)\right], \end{aligned} \tag{10}$$

with a similar expression for country j.

In our analysis, we will assume that country i is willing to sign (or keep) a PTA with country j only if she benefits from the agreement, having the welfare change in (10) be positive. Tariff elimination (or reduction) will increase the gross utility as it expands the consumers' access to commodity varieties as shown below:

$$\triangle_j V^i(\underline{t}^i) = s^j t^i \left\{ \frac{1-\sigma}{(2-\sigma)^2} + \frac{t^i}{4(1-\sigma)}\left[\frac{1}{2} - \left(\frac{\sigma}{2-\sigma}\right)^2 \left(1 - s^{C_i} - \frac{s^j}{2}\right)\right]\right\}, \tag{11}$$

where the last bracketed term in the right side of the equation in (11) is positive for $\sigma \in (0,1)$ as long as $s^{C_i} + s^j/2 > 1/2$. As explained by Furusawa and Konishi

(2007), this sufficiency result reflects the second-best theory that requires the existence of more un-tariffed products than tariffed products for the removal of tariff on the products from country j to have a positive effect on the consumer utility. As one can easily check in (11), a higher s^{C_i} increases $\triangle_j V^i$. Thus, we denote this effect as a *distortion reduction* effect of pre-existing PTAs.

In contrast to having this sufficient condition for a PTA's positive effect on the gross utility, the impact of signing (or keeping) a PTA on the trade surplus is more complex. We can decompose this effect on trade surplus as follows:

$$
\begin{gathered}
\triangle_j\left[X^i(t^{-i}) - M^i(t^i)\right] = \\
= \underbrace{\frac{s^i}{\mu^i}\mu^j\{\frac{1}{2-\sigma} - \frac{t^j}{4(1-\sigma)} + \frac{\sigma}{2(1-\sigma)(2-\sigma)}(1-s^j-\frac{s^i}{2})t^j\}(1-\frac{s^i\sigma}{2-\sigma})t^j}_{\text{Net change in } i\text{'s exports to } j\ (b_1)} \\
\underbrace{-\frac{\mu^j}{\mu^i}s^i s^{\widehat{C}_j}\frac{\sigma(t^j)^2}{2(1-\sigma)(2-\sigma)}(1-\frac{s^i\sigma}{2-\sigma})}_{\text{Concession erosion effect}} \\
\underbrace{-s^j\{\frac{1}{2-\sigma} - \frac{t^i}{4(1-\sigma)} + \frac{\sigma}{2(1-\sigma)(2-\sigma)}(1-s^i-\frac{s^j}{2})t^i\}(1-\frac{s^j\sigma}{2-\sigma})t^i}_{\text{Net change in } i\text{'s imports from } j\ (b_2)} \\
\underbrace{+s^j s^{\widehat{C}_i}\frac{\sigma(t^i)^2}{2(1-\sigma)(2-\sigma)}(1-\frac{s^j\sigma}{2-\sigma})}_{\text{Loss sharing effect}} \\
\underbrace{+\left(1-s^i-s^j\right)\{\frac{1}{2-\sigma} - \frac{t^i}{2(1-\sigma)} + \frac{\sigma}{2(1-\sigma)(2-\sigma)}(1-s^i-\frac{s^j}{2})t^i\}\frac{s^j\sigma t^i}{(2-\sigma)}}_{\text{Net change in } i\text{'s imports from the ROW } (b_3)} \\
+ \underbrace{s^{\widehat{C}_i}\frac{t^i}{2(1-\sigma)}\frac{s^j\sigma t^i}{(2-\sigma)}}_{\text{Competition effect on pre-PTAs}} - \underbrace{(1-s^i-s^j)s^{\widehat{C}_i}\frac{\sigma t^i}{2(1-\sigma)(2-\sigma)}\frac{s^j\sigma t^i}{(2-\sigma)}}_{\text{Competition effect on non-PTAs}}
\end{gathered}
\tag{12}
$$

The decomposed terms in (12) can be classified into two different parts. One is the change in trade surplus that is not affected by pre-existing PTAs and the other is the change affected by the pre-existing PTAs, with the terms of the latter part being multiplied by either $s^{\widehat{C_i}}$ or $s^{\widehat{C_j}}$. The former part is composed of three terms: Net change in country i's exports to country j (b_1, positive effect); net change in country i's imports from country j (b_2, negative effect); and net change in country i's imports from the Rest of the World excluding country j, denoted by the ROW (b_3, positive).

The latter part is also composed of four terms representing different *third-country* effects. The first one is a *concession erosion effect* (negative effect), representing a smaller increase in country i's exports to country j that is caused by country j's having pre-existing PTAs with other (i.e. *third*) countries. The pre-existing preferential market access of these third countries to country j dilutes the potential market access that country i would expect to obtain from signing a PTA with country j in the absence of such pre-existing PTAs of country j. The second one is a *loss sharing effect* (positive effect), representing a smaller increase in country i's imports from country j that is attributable to the country i's pre-existing PTAs with other (i.e. *third*) countries. The pre-existing preferential market access of these third countries to country i dilutes the potential market access that country j would have obtained from signing a PTA with country i in the absence of such pre-existing PTAs.

Comparing with these first two third-country effects that directly affect the trade surplus between country i and country j, the last two third-country effects affect country i's trade surplus with the ROW. The third term, named as *competition effect on pre-existing PTAs* (positive effect), represents a larger decrease in imports from country $i's$ pre-existing PTA member countries than those that would result in the absence of such pre-existing PTAs. If country i and j sign a new PTA, the pre-existing PTA member countries who were enjoying preferential access to country i will face more severe competition from country j than the one they would have faced without such pre-existing PTAs. The last term, denoted by *competition effect on non-PTAs* (negative effect), represents a smaller decrease in imports from country $i's$ non-PTA member countries than those that would result in the absence of such pre-existing PTAs: The existence of pre-existing PTAs of country i makes non-PTA member countries export less to country i countries, thus making the import reduction effect of a new PTA with country j be smaller. Finally, it is straightforward to check that the *competition effect on pre-existing PTAs* dominates *the competition effect on non-PTAs*, having the total third-country effect on the trade surplus of country i with the ROW be positive.

The overall welfare effect of country i from signing or keeping a PTA with country j can be decomposed into two parts: One that would result from the PTA in the absence of pre-existing PTAs with other trading partners (denoted by $\triangle W^i_{no-prePTAs}$, and defined in the following equation) and the other that results from pre-existing PTAs (denoted by the term in the second line in the following equation):

$$\triangle_j W^i = \triangle_j W^i_{no-prePTAs} \left(\equiv \triangle_j V^i + b_1 + b_2 + b_3 - \frac{s^{\widehat{C_i}} s^j \left(t^i\right)^2}{4(1-\sigma)} \left(\frac{\sigma}{2-\sigma}\right)^2 \right) \\ + \frac{\sigma}{2(1-\sigma)(2-\sigma)} \left[s^{\widehat{C_i}} s^j \left(2 - \frac{\left(1-2s^i\right)\sigma}{2(2-\sigma)} \right) (t^i)^2 - \frac{\mu^j}{\mu^i} s^i s^{\widehat{C_j}} (1 - \frac{s^i \sigma}{(2-\sigma)})(t^j)^2 \right] \tag{13}$$

Based on the decomposition in (13), we can obtain the following proposition on the effect of pre-existing PTAs on a country's incentive to sign or keep a PTA:

Proposition 1. *i) Country i's incentive to sign or keep a PTA with country j is positively affected by country i's own pre-existing PTAs, and this positive effect gets larger with an increase in $s^{\widehat{C_i}}$; ii) Country i's incentive to sign or keep a PTA with country j is negatively affected by country j's pre-existing PTAs, and this negative effect gets larger with an increase in $s^{\widehat{C_j}}$*

Pre-existing PTAs affect the welfare change of country i in five different ways, as discussed earlier with regard to the decomposition in (11) and (12). By adding up these pre-existing PTAs' effects into two groups, depending on whether they originate from country i's own pre-existing PTAs or from country j's (i.e., country i's potential PTA partner's) PTAs, (10) shows that pre-existing PTAs of country i and j work in the opposite directions on country i's incentive to sign a PTA with country j. On the one hand, country i's own pre-existing PTA effect, the first term (multiplied by $s^{\widehat{C_i}}$) inside the bracket of the second line in (13), is positive. One the other hand, country j's pre-existing PTA effect, the second term (multiplied by $s^{\widehat{C_j}}$) inside the same bracket, is negative. Country i's incentive to sign a PTA with country j gets stronger with an increase in $s^{\widehat{C_i}}$. This is because the *distortion reduction effect* on the gross utility, the *loss sharing effect* on country i's import from country j, and the total *competition effect on* country i's imports from the ROW are all positive, as discussed above. The partner's pre-existing PTA effect is composed only of the concession erosion effect, which is negative. Thus, country i's incentive to sign a PTA with country j gets weaker with an increase in $s^{\widehat{C_j}}$.

3 Empirical Strategy

3.1 Derivation of Pre-existing PTA Effect Variables

As shown by Proposition 1, a country's own pre-existing PTAs has a positive effect on her incentive to sign (or keep; omitting "keep" from now on to avoid repetition) a PTA and her potential PTA partner's pre-existing PTAs has a negative effect on her incentive to sign it. Empirically testing this hypothesis poses a challenge as a country's own pre-existing PTAs are also the pre-existing PTAs of her PTA partner's partner, generating a positive effect and also a negative one, respectively, on the likelihood of a PTA being signed between them. Based on our theoretical analysis, we construct pre-existing PTA variables that can distinguish the effect of own pre-existing PTAs from that of the partner's pre-existing PTAs.

With regard to the effect of pre-existing PTAs of country i on the likelihood of a PTA being signed between country i and j, the decomposition in (13) implies that

$$\begin{aligned} &\text{Own PTA effect is } s^j s^{\widehat{C}_i}\left\{\frac{\sigma}{2(1-\sigma)(2-\sigma)}(t^i)^2\left[2-\frac{(1-2s^i)\sigma}{2(2-\sigma)}\right]\right\}\text{, and} \\ &\text{Partner's PTA effect is } -\frac{\mu^i}{\mu^j}s^j s^{\widehat{C}_i}\left\{\frac{\sigma}{2(1-\sigma)(2-\sigma)}(t^i)^2\left[1-\frac{s^j\sigma}{(2-\sigma)}\right]\right\} \end{aligned} \tag{14}$$

with the first effect being positive on country i's PTA incentive and the second one being negative on country j's PTA incentive.

In constructing the pre-existing PTA effect variables, we will focus on the terms outside the curly brackets in (14) with the reasons being explained in Appendix 1. First, note that these terms share the common terms $s^j s^{\widehat{C}_i}$ as both effects in (14) reflect the effect of pre-existing PTAs of country i (with the magnitude of such PTAs being measured by $s^{\widehat{C}_i}$) and how such PTAs affect the exports of country j into the market of country i (with that size being measured by s^j). These two channels ($s^{\widehat{C}_i}$ and s^j) together determine the intensity of country i's pre-existing PTAs' influencing the PTA incentives of country i and j.[11]

For country i's incentive to sign a PTA with country j, $s^j s^{\widehat{C}_i}$ measures how her own pre-existing PTAs reduce the export of country j (and the ROW) into her

[11]With regard to the Own PTA effect, it is true that pre-existing PTAs also affect the exports of the ROW, through the *competition effect* shown in (12). Note that the magnitude of the total competition effect is positively affected by the size of s^j as one can check through summation of the last line in (12).

own market compared to a case without such pre-existing PTAs, and we use the following variable to measure this country i's own pre-existing PTA effect in year t:

$$Own\ PTA^{i}_{j,t} \equiv \left(\frac{\text{Import}^{i}_{j}}{\text{Total Import}^{i}}\right) \sum_{k \in \widehat{C}_i} \left(\frac{\text{Import}^{i}_{k}}{\text{Total Import}^{i}}\right) PTA_{ik,t} \tag{15}$$

where "Import$^{i}_{j}$" and "Total Import^{i}" measure country i's imports from country j and country i's total imports, respectively, with "$PTA_{ik,t}$" being a dummy variable for country i and country k ($\neq i$) having a PTA in year t.[12] Thus, we use the share of country j in country i's total imports to measure s^j. s^j represents the mass of differentiated-product-producing firms located in country j (thus, it is not a directly observable variable), and we assume that it is positively correlated with the share of country j in country i's imports. Recall that country i's own pre-existing PTAs affect her incentive to sign a PTA with country j through its effect on the reduction of country i's imports from country j (and the ROW), of which the size is positively affected by the share of country j in country i's imports. Given a positive relationship between s^k and the share of country k in country i's total imports, we can measure $s^{\widehat{C}_i}$ by a variable that sums up $PTA_{ik,t}$ with its weight being the share of country k ($\in \widehat{C}_i$) in country i's total import.

Note that the import share (for example, Import$^{i}_{j}$/Total Import^{i}) variables lack the time subscripts in (15), reflecting that we use their predicted values of the initial-year observations in data. We estimate the predicted values using a simple gravity equation with fixed effects and the log GDP of nations in the dyad as regressors, following Baldwin and Jaimovich (2012). Even though the measure of differentiated-product-producing firms located in country j (i.e., s^j), is a fixed variable by our theoretical model's assumption, the import share variables are likely to be affected by pre-existing FTAs. For example, if country i signs a PTA with a third country ($\neq j$) in year $t-1$, then it will tend to reduce country j's share in country i's total imports in year t. To avoid this kind of simultaneity problem between PTAs and import shares, we use these time-fixed predicted values for import share variables following the literature on pre-existing PTA effects.[13]

[12] We use 5-year lag variables for pre-existing PTAs variables in our main empirical specification, with the reason being provided in the following subsection.

[13] As their "Interdependence" variable to capture the effect of third-country PTAs on signing a new PTA, Egger and Larch (2008) mainly utilize the bilateral distance (thus, time-fixed) weighted measure of pre-existing PTAs for each pair of potential PTA partners.

$-\left(\mu^i/\mu^j\right) s^j s^{\widehat{C}_i}$ measures how country i's pre-existing PTAs reduce the export of country j into country i's market compared to the case without such pre-existing PTAs and how such a reduction affects country j's incentive to sign a PTA with country i. We measure this country j's partner's pre-existing PTA effect (on her incentive to sign a PTA with country i) in year t by

$$Partner's\ PTA^j_{i,t} \equiv -\left(\frac{\mathrm{POP}^i_t}{\mathrm{POP}^j_t}\right)\left(\frac{\mathrm{Import}^i_j}{\mathrm{Total\ Import}^i}\right)\sum_{k\in\widehat{C}_i}\left(\frac{\mathrm{Import}^i_k}{\mathrm{Total\ Import}^i}\right)PTA_{ik,t} \tag{16}$$

where "$(\mathrm{POP}^i_t/\mathrm{POP}^j_t)$" measures country i's population relative to country j's in year t. Compared with the *Own PTA* effect variable for country i in (15), the *Partner's PTA* effect variable for country j only differs from it by having this minus relative population variable being multiplied to it.

As already discussed for the *Own PTA* effect variable, "how country i's pre-existing PTAs reduce the export of country j into country i's market compared to the case without such pre-existing PTAs" can be measured by the variable in (15). So it remains to explain why we need to multiply the minus relative population to this variable to measure how such a reduction in the export of country j into country i's market affects country j's incentive to sign a PTA with country i. The minus sign is easy to understand because a reduction (caused by pre-existing PTA of country i) in the expected increase in country j's export to country i that can result from signing a PTA with country i generates a negative effect on her incentive to sign the PTA. Finally, note that the variable in (15) may under-represent or over-represent this negative effect on country j's incentive to sign a PTA with country i if country j's population is smaller or larger than country i's, respectively. If country j's population gets smaller compared with country i's (holding s^j constant, thus country j's import share in country i's market constant), then the importance of her export to country i in her total welfare will get bigger as the share of her export to country i in her total production is going to get larger. Thus, we multiply "$-(\mathrm{POP}_i/\mathrm{POP}_j)$" to (15) to measure country j's partner's pre-existing PTA effect (on her incentive to sign a PTA with country i) as in (16).

In analyzing the partner's (pre-existing) PTA effect on a CU-member country's incentive to sign a PTA, we pretend that each member of her CU has a pre-existing PTA with her potential PTA partner, as discussed in the introduction: "This is because a CU member country's improved access into her potential PTA partner's market from signing a PTA will be restrained by other CU members' having the same improved access due to the common external trade policy requirement of a

CU for its member countries."[14] For example, for a potential PTA between France and South Korea at year t, the *Partner's PTA* variable for France is created by pretending that Korea already has PTAs with the remaining 27 countries of the European Union (EU) at year t. To check how this way of constructing the *Partner's PTA* variable for a CU member affects its coefficient estimate, we also construct the *Partner's PTA* variable that treats a CU member as no CU member, denoting it by "$Partner's\ PTA^{noCU}$."

3.2 Specification

The primary focus of our empirical analysis is to check whether pre-existing PTAs affect the probability of a PTA being signed between a pair of countries in the ways that our theoretical model predicts, having a positive Own PTA effect and a negative Partner's PTA effect. Because these pre-existing PTA effects have the dynamic nature in affecting the PTA likelihood, we utilize panel data specifications that are similar to the ones of Egger and Larch (2008) and Baldwin and Jaimovich (2012), as explained below.

In utilizing panel data for the analysis of pre-existing PTA effects, Egger and Larch (2008) and Baldwin and Jaimovich (2012) employ two distinct forms of dyadic data, namely, *undirected* and *directed dyadic data*, respectively.[15] In constructing their *Interdependence* variable, a bilateral distance (or predicted trade volume) weighted measure of pre-existing PTAs for each pair of potential PTA partners, Egger and Larch (2008) do not model any "*directed*" relationship between the pair's pre-existing PTAs and each pairing country's incentive to form a PTA. In contrast, the *Contagion* variable of Baldwin and Jaimovich (2012) does reflect a *directed* relationship, with country j's pre-existing PTAs bolstering country i's incentive to sign a PTA with country j. This difference in their pre-existing PTA effect variables induces them to employ these two different forms of data for their analysis.

[14] As emphasized by Ovádek and Willemyns (2019), not all CUs require their members to jointly negotiate a PTA with a non-member country. EAEU (Eurasian Economic Union) and EU are the two CUs that clearly have such a requirement in their agreements, but the members of MERCOSUR and SACU (Southern African Customs Union) also have always jointly negotiated PTAs with non-member countries. In the following analysis, we use the *Partner's PTA* variable constructed by treating the members of these four CUs as having such a requirement. As robustness checks, we also report our regression results using the *Partner's PTA* variables based on both more strict and less strict definitions of CUs in Section 4.3.

[15] See Neumayer and Plümper (2010) for a detailed discussion of dyadic data.

For comparison, we use both undirected and directed dyadic data for our empirical analysis. First, we use the following latent variable specification for the analysis of undirected dyadic data:

$$\begin{gathered}
Y^*_{ij,t} = \beta_0 + \beta_1 \left(OwnPTA^i_{j,t-5} + Own\ PTA^j_{i,t-5}\right) + \beta_2 \left(Partner's\ PTA^i_{j,t-5}\right.\\
\left. + Partner'sPTA^j_{i,t-5}\right) + \beta_3 Interdependence_{ij,t-5} + \beta_4 X_{ij,t-5} + u_{ij} + e_{ij,t},\\
e_{ij,t} \sim G\left(e_{ij,t}\right),\\
Y_{ij,t} = \mathbf{1}\left\{Y^*_{ij,t} > 0\right\} = \begin{cases} 1 & \text{if} Y^*_{ij,t} > 0 \\ 0 & \text{otherwise,} \end{cases}
\end{gathered} \tag{17}$$

where $Y^*_{ij,t}$ is the latent variable for an observed binary variable, $Y_{ij,t}$, which equals 1 if there exists a PTA between country i and j at year t, with $X_{ij,t-5}$, u_{ij}, and $e_{ij,t}$ respectively denoting lagged economic explanatory variables of a dyad of country i and j, an unobserved dyad effect variable, and an error term that is drawn from a symmetric distribution function G, having $i \neq j$. As indicated by the time subscripts, we use 5-year lagged explanatory variables, following Egger and Larch (2008).[16] To reduce simultaneity problems between pre-existing PTA variables and the dependent variable, we include only not-yet-switched dyads, i.e. pairs that could adopt a PTA but have not yet in our estimation, following Baldwin and Jaimovich (2012).[17]

The *Interdependence* variable of Egger and Larch (2008) is constructed to represent an "undirected" relationship between the pair's pre-existing PTAs and their incentive to form a PTA, having one *Interdependence* variable for each pair of (potential) PTA partners. Even though each of our *Own PTA* and *Partner's PTA* variables has a *directed* relationship, we add up *Own PTA* variables for a pair of potential PTA partners as in (17) to create one combined *Own PTA* variable, and conduct the same kind of addition to create one combined *Partner's PTA* variable. We can justify this construction of variables based on the following reasoning: 1) country i and j dyad's switching its status from non-PTA to PTA depends on both countries' incentives to form a PTA; 2) each country's *Own PTA* variable affects each country's incentive to form a PTA in the same manner (i.e., having a common coefficient, β_1) and each country's *Partner's PTA* variable affects each other's

[16]It often takes a couple of years from the start of a PTA negotiation to an initiation of a PTA. For example, the mean and median of this PTA negotiation duration of Republic of Korea with her 17 negotiation partners are 4.29 and 4 years, respectively. As one of robustness checks, we also conduct our empirical analysis based on 3-year lagged explanatory variables.

[17]Thus, a dyad stays in the panel until a PTA being signed, and then it is dropped from the data.

incentive to form a PTA in the same manner (i.e., having a common coefficient, β_2).

One may consider constructing a combined *Contagion* variable in a similar manner. In contrast to *Own PTA* and *Partner's PTA* variables, adding up *Contagion* variables of a pair of potential PTA partners to create one combined variable has the following conceptual problem. If the combined *Contagion* variable has a very high value because each of *Contagion* variables has a high value, reflecting both potential PTA partners' having substantial pre-existing PTAs, then the trade diversion effect of a partner's pre-existing PTAs that would have invoked rent-losing domestic firms to lobby for a PTA will be diluted by a country's own pre-existing PTAs, which in turn mitigates the pair of countries' contagion-based incentive to sign a PTA. This nature of the *Contagion* variable makes us and Baldwin and Jaimovich (2012) rely on directed dyadic data for estimating the effect of this variable on the proliferation of PTAs.

To deal with possible correlations of the dyad's time-variant regressors with the unobserved dyad effect variable, u_{ij}, we use the following two estimation methods. Following Egger and Larch (2008), we first employ the Mundlak-Chamberlain's correlated random effects probit model: This means that we add the average of each explanatory variable across all periods as a separate regressor on the right-hand-side of (17), with $G(\cdot)$ being Normal (0,1), as explained by Chamberlain (1980) and Wooldridge (2002). The second method is the conditional logic estimation, following Baldwin and Jaimovich (2012). As discussed by Chamberlain (1980), we can adopt this method to control for unobserved heterogeneity at the dyad level, obtaining unbiased estimates of the parameters given that $G(\cdot)$ is the logistic distribution. Compared with the Mundlak-Chamberlain's probit model assuming a specific relationship between the unobserved dyad effect variable and other regressors, the conditional logit model has the advantage of not relying on any particular relationship between them, except assuming that $G(\cdot)$ follows the logistic distribution. The limitation of this approach is its applicability only to the sub-sample of dyads that switch their PTA status during the observed period.

Second, we use the following latent variable specification for directed dyadic data:

$$
\begin{aligned}
Y^{i*}_{j,t} = \beta_0 + \beta_1 OwnPTA^{i}_{j,t-5} + \beta_2 Partner'sPTA^{i}_{j,t-5} \\
+\beta_3 Contagion^{i}_{j,t-5} + \beta_4 X_{ij,t-5} + u_i + e_{i,t}, \\
e_{i,t} \sim G\left(e_{i,t}\right), \\
Y^{i}_{j,t} = \mathbf{1}\left\{Y^{i*}_{j,t} > 0\right\} = \begin{cases} 1 & \text{if} Y^{i*}_{j,t} > 0 \\ 0 & \text{otherwise}, \end{cases}
\end{aligned} \tag{18}
$$

where $Y^{i*}_{j,t}$ is the latent variable for an observed binary variable $Y^{i}_{j,t}$, which equals 1 if there exists a PTA between country i and j at year t. $X_{ij,t-5}$ and G are defined as in (17), with u_i and $e_{i,t}$ denoting an unobserved individual (i.e., country i) effect variable and an error term, respectively, having $i \neq j$. Even though a PTA between country i and j requires both countries' approval, the above specification pretends that it only requires country i's to focus on the directed pre-existing PTA effects on country i's incentive to sign a PTA with country j. This feature of directed dyadic data analysis shared by Baldwin and Jaimovich (2012) reflects the motivation of the domino theory for the proliferation of PTAs: Baldwin (1993) intends to explain why country i that was not interested in a PTA with country j changes her position after country j signed up PTAs with other trading partners. For this directed dyadic data analysis, thus we have *directed* pre-existing PTA effect variables in (18), *Own PTA*, *Partners' PTA*, and *Contagion*, instead of undirected ones in (17). Once again, we include only not-yet-switched dyads (i.e., pairs that could adopt a PTA but have not) to reduce simultaneity problems between pre-existing PTA variables and the dependent variable.

Note that both *Partner's PTA* variable and *Contagion* variable in (20) are constructed based on pre-existing PTAs of country j to capture their effects on country i's incentive to sign a PTA with country j, yet predicting opposite signs on their coefficients. Even though both variables utilize the same weights in adding up pre-existing PTAs of country j, namely the predicted import share of country j's PTA partner in country j's total imports, to measure their effects on country j's market, these measures differ from each other in terms of what is multiplied to this weighed measure of pre-existing PTAs. Because the *Contagion* variable is supposed to reflect country i's exporting industry's political incentive to lobby for a PTA with country j to dilute the trade diversion effect caused by country j's pre-existing PTA(s) with other trading partner(s), the multiplied term is the predicted share of country i's export to country j in country i's total exports. In contrast, the *Partner's PTA* variable reflects the degree of concession erosion for country i

caused by pre-existing PTAs of country j, thus the multiplied term is the predicted import share of country i in country j's total imports that is multiplied again by the ratio of country j's population to country i's, as explained in the preceding subsection.

To deal with possible correlations of the dyad's time-variant regressors with the unobserved individual effect variable, u_i in (20), once again we use the two estimation methods, the Mundlak-Chamberlain's correlated random effects probit model and the conditional logic estimation, with the same assumption on G for the corresponding method.

Despite possible correlations between the time-variant regressors and the unobserved dyad effect or individual effect variable (u_{ij} or u_i), we also run a logit estimation of the following specification for undirected dyadic data and the one for directed dyadic data, respectively:

$$\begin{aligned} Pr\left(switch_{ij}\right) = F\Big[\beta_0 + \beta_1 \left(OwnPTA^i_{j,t-5} + Own\ PTA^j_{i,t-5}\right) + \beta_2 \big(Partner's \\ PTA^i_{j,t-5} + Partner'sPTA^j_{i,t-5}\big) + \beta_3 Interdependence_{ij,t-5} + \beta_4 X_{ij,t-5} + \beta_5 X_{ij}\Big] \end{aligned} \tag{19}$$

$$\begin{aligned} Pr\left(switch_{ij}\right) = F\Big(\beta_0 + \beta_1 OwnPTA^i_{j,t-5} + \beta_2 Partner'sPTA^i_{j,t-5} \\ + \beta_3 Contagion^i_{j,t-5} + \beta_4 X_{ij,t-5} + \beta_5 X_{ij}\Big) \end{aligned} \tag{20}$$

where the probability of country i and j dyad's switching its status from non-PTA to PTA, denoted by $Pr(switch_{ij})$, is determined by the logistic cumulative distribution function $F(\cdot)$ of a linear vector of lagged explanatory variables $X_{ij,t-5}$ of the country dyad, as well as our pre-existing PTA effect variables and the *Interdependence* variable in (19) or the *Contagion* variable in (20).[18] Following Baldwin and Jaimovich (2012), we also include dyad-level time-invariant control variables X_{ij}

[18]In addition to lagged explanatory variables of the country dyad, Baldwin and Jaimovich (2012) also include lagged explanatory variable of country i ($X_{i,t-1}$) and lagged explanatory variables (X_t). They employ these additional variables mostly to capture the effects of multilateral trade liberalization on regionalism by including WTO related variables, as well as political determinants such as democracy status. Following Baier and Bergstrand (2004) and Egger and Larch (2008), we focus on the country-pair economic explanatory variables, except the pre-existing PTA effect variables, which in turn facilitates comparing the estimation results from directed dyadic data with those from undirected ones.

(such as geographic distance, common border, and common language), instead of relying on the typical country-pair fixed effect control.[19]

In summary, we employ the Mundlak-Chamberlain's correlated random effects probit model and the conditional logic model for our analysis of both undirected dyadic data and directed dyadic data, based on the specification in (17) and (18), respectively. In addition, we employ the logic estimation based on the specifications in (19) and (20), having six main regression methods in total.

4 Data and Results

4.1 Data

We obtain our data on PTAs from the Database on Economic Integration Agreements (EIAs) of the Kellogg Institute for International Studies and the WTO Regional Trade Agreements (RTAs) Database.[20] The former data set uses the following number-coded classification to reflect the degree of economic integration: 0 for no agreement; 1 for non-reciprocal preferential trade agreements, 2 for preferential trade agreements; 3 for free trade agreements; 4 for customs unions; 5 for common markets; and 6 for economic unions. Note that a non-reciprocal preferential trade agreement (coded by "1") of this classification typically represents the Generalized System of Preferences of a country, providing tariff reduction for least developing countries. The preferential trade agreements coded by "2" are also different from our definition of PTAs as they represent partial preferential trade arrangements that do not eliminate protectionary measures on most of products traded between the members or they adopt such elimination of protectionary measures among a subset of their members.[21] Our dependent variable, a PTA dummy, thus takes a value of 1 if this classification number is equal or greater than 3, and takes 0 otherwise.

[19]As explained by Baldwin and Jaimovich (2012), "estimating panel data model with a limited dependent variable raises inconsistency in the estimation of fixed effects being transmitted to inconsistency in the estimation of parameters."

[20]The Database on EIAs updated in April, 2017 contains the information about ratified RTAs from 1950-2012, thus we update this data using the WTO RTAs Database to include PTAs ratified by the end of 2017. These databases are available on the following websites: https://sites.nd.edu/jeffrey-bergstrand/database-on-economic-integration-agreements; https://rtais.wto.org/UI/PublicMaintainRTAHome.aspx

[21]For example, APTA (Asia Pacific Trade Agreement), LAIA (Latin American Integration Association), and SADC (Southern Africa Development Community) belong to this category.

To include a large set of countries in our analysis, 183 countries, we end up limiting the data period to 25 years from 1993 to 2017 during which PTAs have rapidly proliferated across the world: The cumulative number of PTAs in force was 33 in 1993 and it increased to 293 in 2017. This implies that the maximum number of country-pairs, for which the dependent variable (i.e., PTA dummy) can take zero or one as its value is 16,653 per year, making the maximum possible sample size be 416,325 for our undirected dyadic data analysis and double that number for the analysis of directed dyadic data.[22]

Our main explanatory variables, *Own PTA* and *Partner's PTA*, are constructed as explained in Section 3.1. With regard to other pre-existing PTA variables, *Contagion* and *Interdependence*, we provide a brief explanation of their construction in Section 3.2, referring the exact construction of them to Baldwin and Jaimovich (2012) and Egger and Larch (2008), respectively. The selection of other time-varying explanatory variables largely relies on Baier and Bergstrand (2004), following Egger and Larch (2008), with the time-invariant regressors for logit specifications in (19) and (20) being adopted à la Baldwin and Jaimovich (2012).[23] We define and denote these regressor variables as follows, with the expected signs for their coefficient estimates shown in parentheses:

- *RGDP Sum* (+) = $log(realGDP_{it} + realGDP_{jt})$

- *RGDP Sim* (+) = $log\{1 - [realGDP_{it}/(realGDP_{it} + realGDP_{jt})]^2 - [realGDP_{jt}/(realGDP_{it} + realGDP_{jt})]^2\}$, which measures the similarity of the real GDPs of a country pair.

- *Inverse Distance* (+) measures the log of the inverse distance between the most populated cities of two trading partners.

- *Same Continent* (+) is dummy variable which has the value one if two trading partners are on the same continent, and zero otherwise.

- *Distance from RoW* (+) =$0.5[log(\sum_{k\neq j} distance_{ik}/n-1) + log(\sum_{k\neq i} distance_{kj}/n-1)]$ which measures the average of total distances from ROW of a country pair.

- *K/L Diff* (+) = $|log(realGDP_{it}/population_{it})$ - $log(realGDP_{jt}/population_{jt})|$, which meausres the relative factor endowment difference of a country pair country.

- *K/L Diff sq* (-) = $(K/L\ Diff.)^2$

[22]The actual sample size that we employ in our regression analysis varies depending on the length of lag adopted for regressors as well as the availability of regressor data.

[23]The data sources are as follows: trade data from UN Comtrade; population, GDP, and GDP per capita from World Bank indicators; and Distance, Contiguous, Common language, Colony, and Same Colony from CEPII.

- *K/L Diff from RoW* (-) = $0.5[|log(\sum_{kt\neq it} realGDP_{kt}/\sum_{kt\neq it} population_{kt}) - log(realGDP_{it}/Population_{it})| \quad +|log(\sum_{kt\neq jt} realGDP_{kt}/\sum_{kt\neq jt} population_{kt}) - log(realGDP_{jt}/Population_{jt})|]$, which meausres the average of relative factor endowment difference from ROW of a country pair.

Summary statistics based on the undirected dyadic dataset is presented in Table 1, comparing their mean values with previous studies'. In Table 1, our mean values of *PTA* and *K/L Diff from RoW* are significantly different from those of Egger and Larch (2008). The difference in the mean values of *PTA*s is likely to come from our dropping the observations in the panel after the PTA status is switched from non-PTA to PTA as the mean value of this variable in Baldwin and Jaimovich (2012) who adopt the same dropping rule shows a similar mean value. Including more developing countries in data as well as using more recent data compared with the data of Egger and Larch (2008) using 145 countries during 1955-2005 may explain the difference in mean values for *K/L Diff from RoW*.

Table 1: Summary statistics for undirected dyadic data

	Obs.	Mean	Mean (EL, 2008)	Mean (BJ, 2012)
PTAs	386,219	0.005	0.057	0.007
Own PTA	310,729	0.403	-	
Partner's PTA	309,580	20.457	-	
Partner's PTA^{noCU}	309,580	19.332		
Contagion	310,729	0.581	-	13.185
Interdependence	369,969	0.173	-	33.906
RGDP Sum	333,791	25.487	11.291	49.610
RGDP Sim	333,791	-2.350	-1.952	
Inverse Distance	386,219	-8.840	-8.679	
Distance from RoW	386,219	8.949	8.855	
Same Continent	386,219	0.185	0.237	
K/L Diff	333,471	1.825	1.242	
K/L Diff sq	333,471	4.933	2.344	
K/L Diff from RoW	333,471	6.240	0.992	
Contiguous	386,219	0.009		0.007
Common language	386,219	0.144		0.115
Colony	386,219	0.009		0.029
Same Colony	386,219	0.108		0.051

The mean value of *Interdependence*, 0.173, is less than one because we normalize each row of the weighting matrix of the *Interdependence* variable so that it sums up to unity. With regard to the *Own PTA*, *Partner's PTA*, and *Contagion* variables, however, we do not apply such normalization because it will eliminate some essential information from the weighting matrix such as $\text{Import}^i_j/\text{Total Import}^i$ of $OwnPTA^i_j$, $(\text{POP}^j/\text{POP}^i)(\text{Import}^j_i/\text{Total Import}^j)$ of $Parnter'sPTA^i_j$, and $\text{Export}^i_j/\text{Total Export}^i$ of $Contagion^i_j$.[24] The mean value of the *Partner's PTA* variable is much larger than that of the *Own PTA* or *Contagion* variable because the relative population ratio frequently takes a very large value.[25] We cannot provide or find a reason for why the mean values of *Contagion* and *Interdependence* variables reported by Baldwin and Jaimovich (2012) are much larger than ours.[26]

4.2 Estimation Results

To empirically evaluate the pre-existing PTA effects on the formation of PTAs, we analyze both undirected and directed dyadic data, employing the Mundlak-Chamberlain probit (denoted by M-C Pr.), the conditional logit (denoted by C Logit), and the logit (denoted by Logit) models, as specified in Section 3.2. Table 2 shows the estimation results based on undirected dyadic data, with M-C Probit results in the first four columns and C logit results in the last two columns. While we employ the sum of dyad's *Own PTA*, *Partner's PTA*, and *Contagion* variables specified as in (17) as regressors of this regression analysis, we simply refer to these variables by *Own PTA*, *Partner's PTA*, and *Contagion* in Table 2 and the following discussion of it.

The first column of Table 2 shows the result when we include only our pre-existing PTA effect variables, thus excluding *Interdependence* and *Contagion* from the regression analysis. The coefficient estimates on the *Own PTA* and *Partner's PTA* variables have expected signs and they are statistically significant at the 1% level based on p-statistics. The coefficient estimates on other time-variant regres-

[24]This aspect of these pre-existing PTA variables that does not allow the usual normalization of the corresponding weighting matrix may make comparing the size of coefficient estimates of these variables difficult, which we do not attempt to do in this paper. For an expositional reason, we multiply 1,000 to the original values of these variables.

[25]The mean value of this population ratio is 72.51.

[26]In fact, we expect the mean values of these variables to be less than one if the normalization had been applied to them.

[27]Note that when analyzing with M-C Probit, the Pseudo R^2 was not reported, so it was calculated with the log likelihood of the constant-only model (LLC) and the full model (LLF) of each empirical specification such as R^2 = |LLC-LLF| / LLC.

Table 2: Main results with Undirected Dyadic Data

t-5	Undirected dyadic data					
Method	M-C Pr.	M-C Pr.	M-C Pr.	M-C Pr.	C Logit	C Logit
Own PTA (+)	1.16***	1.15***	0.58***	2.46***	951.55***	2001.12*
Partner's PTA (-)	-0.001***	-	-0.001***	-0.001***	-0.60***	-0.66***
Partner's PTA^{noCU}(-)	-	-0.0002	-	-	-	-
Contagion (+)	-	-	-	-1.93***	-	-961.16
Interdependence (+)	-	-	16.16***	16.38***	8282.80***	8281.91***
RGDP Sum (+)	1.03***	1.03***	-0.64***	-0.50**	26.37***	26.38***
RGDP Sim (+)	1.22***	1.28***	0.65***	0.62***	-0.32	-0.33
K/L Diff. (+)	0.19**	0.21**	0.20*	0.21**	-0.30	-0.30
K/L Diff. sq (-)	0.03**	0.03**	0.04***	0.05***	-0.14	-0.14
K/L Diff. from RoW (-)	1.53***	1.56***	0.90***	0.99***	6.55*	6.54*
obs.	237,966	237,966	237,966	237,966	15,193	15,193
(Pseudo) R^2[27]	0.120	0.120	0.165	0.177	0.957	0.958

Note: *. **, and *** denote significance at 10%, 5%, and 1% respectively.

sors mostly have expected signs that are statistically significant at the 1% or 5% levels, except for the ones on *K/L Diff sq* and *K/L Diff from RoW*.[28] A positively signed coefficient estimate on *K/L Diff from RoW* also occurs in the analysis of Egger and Larch (2008), and this common difference from the corresponding result of Baier and Bergstrand (2004) might be caused by the difference in the empirical methodologies adopted across these studies.[29]

The second column shows the result when we only replace the *Partner's PTA* variable with $Partner's\ PTA^{noCU}$. Then, the coefficient estimate on this variable gets smaller by half in absolute value and becomes statistically insignificant. Thus, not-pretending "that each member of her CU has a pre-existing PTA with her potential PTA partner" leads to an under-estimation of the *Partner's PTA* effect: a CU member country is less likely to sign a new PTA than a non-member country because each member's gain from signing a new PTA will be

[28]Having a positive sign on the coefficient estimate of *K/L Diff sq* seems to be less of a problem compared with having a positive sign on the one of *K/L Diff from RoW* because we have an expected sign on the coefficient estimate on *K/L Diff*: It simply implies the positive effect of *K/L Diff* on signing a PTA does not decrease as *K/L Diff* increases. Furthermore, the coefficient estimate of *K/L Diff sq* changes its sign from positive to negative in the C-Logit and Logit estimations as shown in Table 2, 3, and 4.

[29]The coefficient estimate on *K/L Diff from RoW* changes its sign from positive to negative when we adopt the logit estimation, as shown in Table 4.

diluted by other members' having the same preferential access to the new PTA partner's market! While not reported in Table 2, losing the statistically significance on the coefficient estimate of *Partner's PTA* variable when we change it with $Partner's\ PTA^{noCU}$continues to occur in the M-C Probit that includes the *Interdependence* variable and in the C Logit without the *Interdependence* variable. This result on CU members' PTA formation incentive provides an indirect empirical support for the previous theoretical works on the CU's less preferred aspect of unanimity-based membership expansion rule in attaining global free trade, such as Yi (1996) and Furusawa and Konishi (2007).[30]

Introducing the *Interdependence* variable into the regression does not affect the sign and the significance of coefficient estimates on our pre-existing PTA effect variables, as shown in the third column of Table 2. Thus, Own and Partners' pre-existing PTA effects influence the formation of PTAs according to the prediction of our theoretical model, together with the *Interdependence* variable, for which there is no theoretical ground to explain how it affects the formation of PTAs. Other time-varying regressors continue to have the same signed coefficient estimates as those in the first and the second columns, except the one on *RGDP Sum* variable changing its sign, for which we have no persuasive explanation.[31]

With regard to using the sum of dyad's *Contagion* variables in our analysis of undirected dyadic data, Section 3.2 provides a discussion of why constructing such a variable is subject to a conceptual problem, possibly invalidating the use of the *Contagion* variable in our regression analysis. Despite this concern, we may still include this variable together with other pre-existing PTA effect variables in the M-C Probit regression, reporting the associated results in the fourth column in Table 2. While the inclusion of the *Contagion* variable does not affect the signs (and their statistical significance) of the coefficient estimates of other regressors, this variable has a statistically significant negative coefficient estimate, a result that may seem to contradict the domino theory's prediction: An increase in a potential PTA partner's pre-existing PTAs seems to discourage potential PTA partners from signing a new PTA according to the regression analysis, rather than encouraging it as the theory predicts. However, it is important to note that this result is not necessarily empirical evidence against the domino theory. The sum of the dyad's *Contagion* variables may take a higher value when the number of pre-

[30]We provide a detailed discussion of these theoretical works and their relation with this empirical result in the introduction section.

[31]The coefficient estimate on *RGDP Sum* variable takes a negative value only when we include the *Interdependence* variable in the M-C Probit regressions, but takes a statistically significant positive value in all other regressions.

existing PTAs of both potential PTA partners increases, possibly in a symmetric way. Such an increase in both partners' pre-existing PTAs may nullify the trade diversion effect of each other's pre-existing PTAs that invokes the political lobby for a new PTA between them, thus their domino effects, as discussed in 3.2.

The results from the C logit in Table 2 largely produce the same results as the corresponding ones from the M-C Probit. The results on pre-existing PTA variables in the fifth column are qualitatively the same as those in the third column, re-confirming our prediction on these variables specified in (17). The coefficient estimate on *RGDP Sum* now has the expected positive sign that is statistically significant, and the coefficient estimates on RGDP Sim, K/L Diff. and K/L Diff. sq change their signs but they are no longer statistically significant. A possible cause for this loss of statistical significance is the large reduction in the sample size because the C logit can be run only on the sub-sample of dyads that switch FTA status during the observed period, having the observation reduced from 237,966 to 15,193.[32] Including *Contagion* as a regressor in the C logit generates once again the same outcomes as the inclusion of this variable in the M-C Probit does, having a negative coefficient estimate on it, but this estimate is no longer statistically significant, as shown in the sixth column of Table 2.

Table 3 shows the estimation results based on directed dyadic data, with the M-C Probit results in the first four columns and the C logit results in the last two columns. Recall that the pre-existing PTA effect variables in these estimations are *directed* dyadic variables as specified in (18), except the *Interdependence* variable. The dependent variable (i.e., PTA dummy) enters the regressions twice for each dyad, which in turn doubles the sample size comparing with the corresponding regressions in Table 2.

The first two columns of Table 3 show the results when we include only our pre-existing PTA effect variables with *Partner's PTA* variable in the first one and $Partner's\ PTA^{noCU}$ variable in the second. The estimation results in these columns are qualitatively identical to those in the first two columns in Table 2, except the coefficient estimate on $Partner's\ PTA^{noCU}$ carrying a positive sign that is statistically significant at the 10% level, instead of carrying a statistically insignificant negative sign. Thus, replacing *Partner's PTA* with $Partner's\ PTA^{noCU}$ once again makes the coefficient estimate on this variable no longer carry a statistically significant negative sign, implying that "a member country of CU is less likely to sign a new PTA than a non-member country because each member's gain from signing

[32]A lot of dyads that are eliminated from the sample are the ones that includes least developed countries as they typically remain not having any PTA during the whole sample period. This must have significantly reduced the variations in *RGDP Sim*, *K/L Diff*, *K/L Diff sq*, and *K/L Diff from RoW*.

Table 3: Main results with Directed Dyadic Data

t-5	Directed dyadic data					
Method	M-C P.	M-C P.	M-C P.	M-C P.	C Logit	C Logit
Own PTA (+)	1.35***	1.23***	1.30***	0.55***	116.65***	930.46***
Partner's PTA (-)	-0.001**	-	-0.001***	-0.001***	0.03	-0.65***
Partner's PTA^{noCU} (-)	-	0.0005*	-	-	-	-
Contagion (+)	-	-	1.89***	0.42***	137.22***	890.58***
Interdependence (+)	-	-	-	16.31***	-	8300.27***
RGDP Sum (+)	1.81***	1.59***	1.34***	-0.64***	45.11***	26.36***
RGDP Sim (+)	1.65***	1.49***	1.53***	0.67***	15.12***	-0.27
K/L Diff (+)	0.23***	0.22***	0.22***	0.20***	3.22***	-0.28
K/L Diff sq (-)	0.02***	0.02***	0.03***	0.04***	-0.33***	-0.14
K/L Diff from RoW (-)	1.96***	1.77***	1.88***	0.88***	7.84***	6.56**
obs.	475.932	475.932	475,932	475,932	30,386	30,386
(Pseudo) R^2	0.102	0.102	0.116	0.163	0.842	0.957

Note: *. **, and ***denote significance at 10%, 5%, and 1% respectively.

a new PTA will be diluted by other members' having the same preferential access to the new PTA partner's market!"[33]

Including the *Contagion* variable in the regression does not affect the qualitative results on other regressors, as shown in the third column of Table 3. The coefficient estimate on the *Contagion* variable has a statistically significant positive sign, following the domino theory's prediction. Because we can label the positive coefficient estimate on *Own PTA* as a *snowballing effect* of pre-existing PTAs, as discussed in the introduction, this result in the third column demonstrates the empirical existence of snowballing alongside domino effects in the recent proliferation of PTAs. Despite the similarity between *Partner's PTA* and *Contagion* in the construction of these variables, it is also worthwhile to note that they carry opposite signs on their coefficient estimates, as predicted by our and domino theories of pre-existing PTA effects.

Adding the *Interdependence* variable into our baseline specification of (18) affects neither signs nor statistical significance of the coefficient estimates of other pre-existing PTA variables, as shown in the fourth column of Table 3. The coefficient estimate on *the Interdependence* variable has a statistically significant positive

[33] While not reported in Table 3, replacing *Partner's PTA* with *Partner's* PTA^{noCU} in our main specification of (18) also makes the coefficient estimate on this variable no longer carry a statistically significant negative sign in the M-C Probit.

sign.

The fifth and sixth columns of this table show that the results on these pre-existing PTA variables from the C Logit based on the specification of (18) or based on the one with the *Interdependence* variable stays the same as the corresponding results in the third and forth columns, respectively, except the coefficient estimate on the *Partner's PTA* variable having a statistically insignificant positive sign in the fifth column. The elimination of dyads that do not switch their PTA status from the data set in the C Logit may have contributed to this loss of statistical significance of the coefficient estimate on the *Partner's PTA* variable: The Partner's PTA effect may have kept at least some part of dyads in data from switching their PTA status from non-PTA into PTA, thus elimination of such data may weaken the Partner's PTA effect in the C Logit.

With regard to changes in the results on other time-variant regressors when we add the *Interdependence* variable or change the estimation method from the M-C Probit to the C Logit, we can summarize them as follows. As already mentioned in Footnote 31 about Table 2, the coefficient estimate on *RGDP Sum* variable takes a negative value only when we include the *Interdependence* variable in the M-C Probit, but takes a statistically significant positive value in all other regressions. In the C Logit, the coefficient estimate on *K/L Diff sq* changes its sign into a negative one, but it loses its statistical significance when we include the *Interdependence* variable. The loss of statistical significance when we include the *Interdependence* variable in the C Logit also happens to the coefficient estimates on *RGDP Sim* and *K/L Diff*, as shown in the sixth column of Table 3.[34]

Table 4 reports the results from running the Logit estimation: The first column to the third one report the results based on the specification of (19) for undirected dyadic data, and the forth column to the sixth one report the results based on the specification of (20) for directed dyadic data.

With regard to the pre-existing PTA effect variables for the analysis of undirected dyadic data, the coefficient estimates on the *Own PTA* and *Partner's PTA* variables have expected signs that are statistically significant at the 1% level, with or without *Interdependence* variable, as shown in the first two columns of Table 4.[35] The *Interdependence* variable, however, lose its explanatory power in the Logit estimation, having a statistically insignificant positive sign on its coefficient estimate.

[34]Recall that the same loss of statistical significance occurs in the C Logit with undirected dyadic data when we include the *Interdependence* variable.

[35]While the *Own PTA*, *Partner's PTA*, and *Contagion* variables are denoted without referring possible summation of these variables for each dyad in Table 4, we do use such summed variables for the analysis of undirected dyadic data.

Table 4: Logit estimation results

t-5	Undirected dyadic data			Directed dyadic data		
Method	Logit	Logit	Logit	Logit	Logit	Logit
Own PTA (+)	0.76***	0.76***	0.10	0.41***	0.79***	0.78***
Partner's PTA (-)	-0.002***	-0.002***	-0.002***	-0.001***	-0.003***	-0.003***
Contagion (+)	-	-	0.69***	-	0.68***	0.67***
Interdependence (+)	-	0.10	0.09	-	-	0.16
GDP Sum (+)	0.14***	0.14***	0.18***	0.23***	0.17***	0.17***
GDP Sim (+)	0.18***	0.18***	0.24***	0.25***	0.23***	0.23***
Inverse Distance (+)	0.77***	0.76***	0.78***	0.83***	0.78***	0.77***
Distance from RoW (+)	0.94***	0.94***	1.28***	0.08	1.04***	1.05***
Same Continent (+)	0.42***	0.42***	0.42***	0.21***	0.40***	0.40***
K/L Diff (+)	0.11***	0.11***	0.07***	0.10***	0.09***	0.09***
K/L Diff sq (-)	-0.03***	-0.03***	-0.02***	-0.03***	-0.03***	-0.03***
K/L Diff from RoW (-)	-0.22***	-0.21***	-0.19***	-0.33***	-0.21***	-0.20***
Contiguous	0.21	0.22	0.20	0.34***	0.21**	0.22**
Common language	0.54***	0.54***	0.53***	0.54***	0.54***	0.54***
Colony	-0.04	-0.05	-0.15	0.07	-0.10	-0.09
Same Colony	-0.31***	-0.31***	-0.25**	-0.34***	-0.29***	-0.28***
obs.	237,966	237,966	237,966	475,932	475,932	475,932
(Pseudo) R^2	0.145	0.145	0.147	0.133	0.145	0.145

Note: *. **, and *** denote significance at 10%, 5%, and 1% respectively.

If we add the *Contagion* variable in the analysis of undirected dyad data despite the conceptual issue associated with it, then the positive coefficient estimate on *Own PTA* loses its statistical significance. The third column also shows that the *Contagion* variable carries a statistically significant positive sign on its coefficient estimate, a result that contrasts to the ones found in Table 2.

The coefficient estimates on other time-varying regressors, originated from the analysis of Baier and Bergstrand (2004), have all the expected signs that are statistically significant. Among the four dyad-level time-invariant control variables, the *Contiguous* and *Colony* variables have no explanatory power in the regression analysis of undirected dyadic data. These results on non pre-existing PTA effects variables stay qualitatively the same regardless of whether we analyze undirected or directed dyadic data, expect the positive coefficient estimate on the *Contiguous* variable becoming statistically significant in the latter analysis.

The analysis of directed dyadic data based on the Logit in Table 4 generates the results that are very similar to the ones based on the M-C Probit or C Logit estimations reported in Table 3. The coefficient estimates on the *Own PTA* and *Partner's PTA* variables have expected signs that are statistically significant, with or without the *Contagion* variable that also carries a statistically significant expected sign, as shown in the fourth and the fifth columns of Table 4. Thus, this Logit estimation also provides empirical support for the existence of *snowballing alongside domino effects* in the recent proliferation of PTAs!

Adding the *Interdependence* variable to the specification of (20) does not affect the effects of all other variables on the formation of PTAs, as shown in the sixth column in Table 4. The coefficient estimate on the *Interdependence* variable has a statistically insignificant positive sign, a result that Baldwin and Jaimovich (2012) also find in their analysis based on the logit estimation.

While not reported in Table 4, replacing *Partner's PTA* with $Partner's\ PTA^{noCU}$ does not affect the statistically significant negative sign on this variable, a result that contrasts to what we find in the M-C Probit and C Logit in Table 2 and 3. A possible reason for this discrepancy is the lack of considering (or under-representing) the unobserved dyad or individual effect in the Logit: Even $Partner's\ PTA^{noCU}$ turns out to have a statistically significant negative effect on PTA proliferation in the Logit because this variable is possibly correlated with the unobserved effect that generates such a negative effect on signing new PTAs.

4.3 Robustness Checks

In this subsection, we conduct several robustness checks.

4.3.1 3-year Lagged variables

First, we re-run the main regressions of Section 4.2 using 3-year lagged variables instead of 5-year lagged variables. Table 5 show the results from this analysis, having the results from the three different regressions using undirected dyadic data in the first three columns and the ones using directed dyadic data in the last three columns.

With regard to the analysis of undirected dyadic data, we only report the results from the main specifications of (17) and (19) in Section 3.2 as the results in other specifications remain qualitatively the same for pre-existing PTA effect variables. For example, replacing the *Partner's PTA* variable with $Partner's\ PTA^{noCU}$ variable in the first two columns yields the same results as the ones from the corresponding regressions with 5-year lagged variables: The coefficient estimate on this

Table 5: Estimation results with 3-year lagged variables

t-3	Undirected dyadic data			Directed dyadic data		
Method	M-C P.	C Logit	Logit	M-C P.	C Logit	Logit
Own PTA (+)	0.48***	10.97***	0.80***	1.03***	32.56***	0.84***
Partner's PTA (-)	-0.001***	-0.01***	-0.002***	-0.001***	-0.03***	-0.003***
Contagion (+)	-			-	40.05***	0.72***
Interdependence (+)	13.00***	399.54***	0.13	0.93***		
GDP Sum (+)	0.43**	47.271***	0.11***	3.38***	51.75***	0.13***
GDP Sim (+)	1.27***	23.61***	0.16***	3.12***	29.69***	0.22***
Inverse Distance (+)	-	-	0.77***	-	-	0.79***
Distance from RoW (+)	-	-	0.91***	-	-	1.03***
Same Continent (+)	-	-	0.56***	-	-	0.54***
K/L Diff (+)	0.41**	-6.30**	0.37***	0.31**	-10.18***	0.36***
K/L Diff sq (-)	-0.07*	-1.03	-0.09***	-0.05	0.12	-0.09***
K/L Diff from RoW (-)	2.26***	34.12***	-0.20***	5.06***	38.26***	-0.20***
Contiguous	-	-	0.14	-	-	0.13
Common language	-	-	0.51***	-	-	0.50***
Colony	-	-	0.006	-	-	-0.06
Same Colony	-	-	-0.30***	-	-	-0.28***
obs.	269,800	18,276	269,800	539,600	36,552	539,600
(Pseudo) R^2	0.167	0.909	0.158	0.138	0.871	0.158

Note: *. **, and ***denote significance at 10%, 5%, and 1% respectively.

variable becomes statistically insignificant. Our pre-existing PTA effect variables, *Own PTA* and *Partner's PTA*, continue to carry the coefficient estimates with expected signs that are statistically significant in all three regression methods. Once again, the coefficient estimate on the *Interdependence* variable loses its statistical insignificance in the Logit estimation.

The results on the coefficient estimates on other time-varying regressors remain qualitatively the same for most of the regressors. The coefficient estimate on *K/L Diff sq* in the M-C Probit changes its sign into a negative one that is statistically significant at the 10% level, as expected by Baier and Bergstrand (2004). In the C Logit, the coefficient estimate on *K/L Diff* changes its sign into a negative one at the 5% significance level. As in the case of 5-year lagged variables, the coefficient estimates on these time-varying regressors originated from Baier and Bergstrand (2004) have all the expected signs that are statistically significant in the Logit. Among the four dyad-level time-invariant control variables in the

Logit, once again only *Common Language* and *Same Colony* variables have statistically significant coefficient estimates.

The last three columns in Table 5 show the results from analyzing directed dyadic data based on the main specifications of in (18) and (20).[36] All pre-existing PTA effect variables, *Own PTA*, *Partner's PTA*, and *Contagion*, continue to have the coefficient estimates with expected signs that are statistically significant in all three regression methods.[37]

Again, the results on the coefficient estimates on other time-varying regressors remain qualitatively the same for most of the regressors, with the changes on *K/L Diff* and *K/L Diff sq* variables across different specifications being similar to the ones described above for the analysis of undirected dyadic data in Table 5. The coefficient estimates on these time-varying regressors originated from Baier and Bergstrand (2004) continue to have all the expected signs that are statistically significant in the Logit. Among the four dyad-level time-invariant control variables in the Logit, now the *Contiguous* variable loses its explanatory power.

4.3.2 Controlling Preference towards PTAs

The next robustness check is about controlling for the existence of a country's preference toward PTAs, such as her preference toward liberal (or equivalently, protectionary) trade policy. The use of an average applied MFN tariff of a country, however, is subject to a simultaneity problem with her pre-existing PTAs, as discussed in Appendix 1. To deal with this problem, we use the average bound tariff variable, denoted by *Bound Tariff*. Bound tariffs are negotiated through the Uruguay round negotiations or through the access negotiations into the WTO for its new members. Because the once-negotiated bound tariffs are fixed, thus the use of average bound tariffs is more desirable than the use of average MFN tariffs, considering the simultaneity problem.

The limitation of *Bound Tariff* is that it is not a time-varying variable so that we cannot include it as a regressor when the regression method is the M-C Probit or

[36]Once again, the results in other specifications remain qualitatively the same for the pre-existing PTA effect variables. For example, replacing the *Partner's PTA* variable with *Partner's* PTA^{noCU} variable in the fourth and fifth columns yields the same results as the ones from the corresponding regressions with 5-year lagged variables, with the coefficient estimate losing its statistical significance in the M-C Probit, but no in the C Logit. Including the *Interdependence* as an additional regressor does not change any of the results from the regressions with 5-year lagged regressors.

[37]In fact, the negative coefficient estimate on the *Partner's PTA* variable gains its statistical significance in the C Logit with 3-year lagged variables, as shown in Table 5.

C Logit.[38] For these regression methods, thus we use the *PTA coverage* variable, a country specific weighted average of her pre-existing PTAs. To be precise, it is the percentage of total exports covered by pre-existing PTAs, with the export shares being the predicted values for the initial year of our data from the gravity regression that we use for constructing the *Own PTA* and *Partner's PTA* variables. As one can expect, the *PTA coverage* and *Own PTA* variables are highly correlated, possibly causing the multicolinearity in the regression analysis when both variables are included as regressors.[39] Despite this potential issue of using this *PTA coverage* variable as a control for the preference toward signing new PTAs, Table 6 shows the results from including this variable for the M-C Probit or C Logit estimations.[40]

Table 6 shows the results from running regressions of the same specifications as in Table 5 (except utilizing the 5 year-lagged regressor variables), plus *PTA coverage* in the M-C Probit and C Logit and *Bound Tariff* in the Logit.

With regard to the analysis of undirected dyadic data, the result from the M-C Probit in the first column of Table 6 shows that the inclusion of the *PTA coverage* variable does not qualitatively affect the coefficient estimates of regressors of the specification (17), except generating a coefficient estimate on *Own PTA* that is more than three times larger than before. Somewhat surprisingly, the coefficient estimate on *PTA coverage* is negative at the 1% significance level. While these coefficient estimates are subject to the multicolinearity issue, a possible explanation for a negative coefficient on *PTA coverage* is given as follows: If we properly measure the effect of *PTA coverage* on signing a new PTA after controlling its relationship with the unobserved dyad level effect variable (through its average value that reflects the liberal preference of a dyad), an increase in *PTA coverage* may in fact generate a negative effect on a dyad's signing a PTA.

The result on the second column of Table 6 shows that the inclusion of the

[38]For the M-C Probit, including a constant variable as a regressor implicitly assumes that the unobserved dyad (or individual) effect variable has no constant term in its statistical relationship with other time-varying regressor variables, an assumption that has no theoretical ground for our PTA formation analysis.

[39]The correlation between *Own PTA* and *PTA coverage* is 0.996 for both undirected and directed data. VIF (Variance Inflation Factor) is 41.49 for *PTA coverage* and 22.79 for *Own PTA* for the case of undirected dyadic data; VIF is 15.70 for *PTA coverage* and 16.40 for *Own PTA* for the case of directed dyadic data. If VIF is greater than 10, then the multicollinearity is considered high.

[40]Baldwin and Jaimovich (2012) use this *PTA coverage* variable in their regression analysis as a control for a country's preference toward PTAs, denoting it by *FTA coverage* variable. Thus, inclusion of this variable also plays the role of checking that our *Own PTA* variable plays a role that is different from *FTA coverage* of Baldwin and Jaimovich (2012), even though the multicolinearity issue may undermines it.

Table 6: Estimation results including *Bound_Tariff* or *PTA_Coverage*

t-5	Undirected dyadic data			Directed dyadic data		
Method	M-C P.	C Logit	Logit	M-C P.	C Logit	Logit
Own PTA (+)	2.17***	2616.34***	0.89***	3.15***	177.61*	0.85***
Partner's PTA (-)	-0.001***	-0.68***	-0.004***	-0.001***	0.03	-0.005***
Contagion (+)	-	-	-	0.76***	137.64***	0.87***
Interdependence (+)	19.85***	8279.78***	1.41***	-	-	-
PTA coverage (+)	-10.63***	-10904.11***	-	-16.80***	-438.36	-
Bound Tariff (-)	-	-	-0.53***	-	-	-0.48***
RGDP Sum (+)	-0.34*	26.38***	-0.004	1.38***	45.21***	0.09***
RGDP Sim (+)	1.06***	-0.34	0.10***	1.14***	15.14***	0.18***
Inverse Distance (+)	-	-	0.29***	-	-	0.56***
Distance from RoW (+)	-	-	0.97**	-	-	1.29***
Same Continent (+)	-	-	0.89***	-	-	0.62***
K/L Diff (+)	0.30**	-0.30	0.26***	0.16*****	3.25**	0.16***
K/L Diff sq (-)	0.05***	-0.14	-0.04***	0.02***	-0.33***	-0.03***
K/L Diff from RoW (-)	1.14***	6.54*	-0.25***	1.42***	7.86***	-0.22***
Contiguous	-	-	0.25	-	-	0.26
Common language	-	-	0.17	-	-	0.26***
Colony	-	-	0.01	-	-	-0.05
Same Colony	-	-	-0.60***	-	-	-0.45***
obs.	237,966	15,193	105,890	475,932	30,386	295,725
(Pseudo) R^2	0.172	0.958	0.159	0.302	0.842	0.145

Note: *. **, and ***denote significance at 10%, 5%, and 1% respectively.

PTA Coverage variable in the C Logit hardly affects the coefficient estimates of the regressors of the specification (17) in Table 2, except having a more than two and half times larger coefficient estimate on *Own PTA* than before. Once again, the coefficient estimate on *PTA coverage* is negative and statistically significant.

The inclusion of *Bound Tariff* in the Logit for its main specification of (19) does not affect the regression result qualitatively, as shown in the third column of Table 6, except making the coefficient estimate on the *Interdependence* variable a statistically significant positive one. The coefficient estimate on *Bound Tariff* carries a statistically significant expected sign (negative).

The last three columns of Table 6 show the results from including our control variable for the possible existence of a country's preference toward signing a new PTA in our analysis of directed dyadic data. Inclusion this control variable does not qualitatively affect the coefficient estimates on existing regressors, except the one on the *Own PTA* variable in the C Logit. The coefficient estimate on *Own PTA* variable continues to be positive but it is statistically significant only at the 10% level. The coefficient estimate on *PTA coverage* is negative but not statistically significant, possibly reflecting the multicolinearity issue associated with introducing *PTA coverage* into this regression. Given the strong correlation between the *PTA coverage* and *Own PTA* variables, one may wonder which variable is a proper variable to include in the regression analysis. In addition to our theoretical justification for *Own PTA*, the Lasso (Least absolute shrinkage and selection operator) analysis for the choice of proper regressors strongly prefers *Own PTA* over *PTA coverage.*[41]

In summary, the results in Table 6 suggest that our main results are mostly robust against including this control variable for PTA preference.

4.3.3 Less Stringent Definition of CU

As discussed in footnote 14 of Section 3.1, not all CUs require their members to jointly negotiate a new PTA in their agreements. In addition to the EAEU and the EU that clearly require such a joint negotiation, we also include MERCOSUR and SACU in our definition of CUs for the preceding analysis (i.e., constructing the *Partner's PTA* variable by assuming that only these four CUs negotiate a new PTA jointly), as the members of these two CUs have always jointly negotiated PTAs with non-member countries. In this subsection, we conduct a robustness check on whether our result changes if we use either a more stringent or a less stringent definition of a CU.

A more stringent definition of a CU would consider only the EAEU and the EU as CUs in constructing the *Partner's PTA* variable. Because the regression analysis based on this more stringent definition of CU generates practically the same results as the preceding analysis, we focus on the regression analysis based on a less stringent definition of a CU. Following the CU definition of Database on EIAs

[41]The Lasso is a regression analysis method for variable selection and regularization to improve the prediction accuracy and interpretability by forcing the sum of the absolute value of the regression coefficients to be less than a fixed value. As shown by Tibshirani (1996), employing the Lasso makes certain coefficient estimates be zero, excluding the corresponding regressors from affecting prediction. In the Lasso analysis of our data with all of our regressor variables, *PTA coverage* ends up being excluded in both the undirected dyadic data and directed dyadic data analysis. In the Lasso analysis of undirected dyadic data, the *Contagion* and *Colony* variables are also excluded, and the *Own PTA* variable is selected first, followed by the *Interdependence*, *Inverse Distance*, and *Bound Tariff* variables.

Table 7: Estimation results with a Less Stringent Definition of CU

t-5	Undirected dyadic data			Directed dyadic data		
Method	M-C P.	C Logit	Logit	M-C P.	C Logit	Logit
Own PTA (+)	0.57***	366.97	0.77***	1.34***	116.61***	0.80***
Partner's PTA (-)	-0.001***	-0.005	-0.003***	-0.002***	0.03	-0.003***
Contagion (+)	-	-	-	1.20***	137.17***	0.68***
Interdependence (+)	15.96***	8317.56***	0.13			
GDP Sum (+)	-0.64***	26.35***	0.14***	1.36***	45.10***	0.16***
GDP Sim (+)	0.64***	-0.21	0.15***	1.54***	15.13***	0.22***
Inverse Distance (+)	-	-	0.76***	-	-	0.78***
Distance from RoW (+)	-	-	0.98***	-	-	1.07***
Same Continent (+)	-	-	0.43***	-	-	0.41***
K/L Diff (+)	0.20*	-0.27	0.10***	0.22***	3.23***	0.09***
K/L Diff sq (-)	0.04***	-1.14	-0.03***	0.03***	-0.33***	-0.03***
K/L Diff from RoW (-)	0.89***	6.57*	-0.21***	1.89***	7.84***	-0.21***
Contiguous	-	-	0.22	-	-	0.21*
Common language	-	-	0.55***	-	-	0.54***
Colony	-	-	-0.05	-	-	-0.10
Same Colony	-	-	-0.30***	-	-	-0.28***
obs.	237,966	15,193	237,966	475,932	30,386	475,932
(Pseudo) R^2	0.165	0.957	0.146	0.116	0.842	0.145

Note: *. **, and *** denote significance at 10%, 5%, and 1% respectively.

described in Section 4.1, we classify PTAs as CUs if their classification number is equal to or greater than 4.[42] Then, Table 7 shows the results from running regressions of the same specifications as in Table 5, except utilizing the 5-year lagged regressors.

Comparison of the results in Table 7 with the corresponding results of the preceding analysis in Section 4.2 show that employing the less stringent definition of CU does not qualitatively affect the results, except the one from the C Logit using undirected dyadic data. The second column in Table 7 demonstrates that the coefficient estimates on the *Own PTA* and *Partner's PTA* variables are no longer

[42]The EAEU, EU, CACM (Central American Common Market), CAN (Andean Community), CARICOM (Caribbean Community), CEMAC (Central African Economic and Monetary Community), COMESA (Common Market for Eastern and Southern Africa), ECOWAS (Economic Community of Western African States), GCC (Gulf Cooperation Council), MERCOSUR, and SACU are such customs unions,

statistically significant even though they still carry the expected signs. Treating all CUs as if they require joint PTA negotiations with non-members in constructing the *Partner's PTA* variable (even though most of such CUs do not impose such requirement on their members in practice) may have contributed to the loss of statistical significance of the coefficient estimate on this variable.

However, we cannot strongly support such an interpretation of the result on the second column because the preceding results stay qualitatively the same in all other specifications. The fact that the results from the more stringent definition of a CU as well as those from the less stringent one remain qualitatively the same, except for this one specification, may imply that the negative effect of CUs on the proliferation of PTAs (through the negative *Partner's PTA* effect) largely comes from the EU, whose members are clearly required to conduct joint PTA negotiations.[43]

5 Calibration Analysis

In this section, we conduct some calibration exercises based on our theoretical model. The primary goal of calibrating our model is to evaluate how well our model can perform in mimicking the actual formation of PTAs during the period of 1993-2017. To achieve this goal, we first introduces ice-berg type trade costs (in addition to tariffs already included) into our model because such trade costs affect the formation of PTAs significantly as shown by many empirical studies on it, including ours. Then, we construct the variables of our model based on actual data. This second stage is not trivial as one cannot find data that directly fits some of our key variables. For example, we need to convert the ad valorem tariff and trade cost data into the specific tariff and trade cost variables of our model. Because such a conversion is sensitive to the substitutability parameter of our quadratic utility function, we end up constructing these variables using a method of numerical approximation based on our model. The last stage is to calibrate our model parameters so that our model's prediction of PTA formation mimics the actual formation of PTAs as closely as possible.

[43]The EAEU is ratified in 2015 so that its effect on the *Partner's PTA* variable is very limited in our analysis of the period from 1993 to 2017. The EU is also composed of up to 28 countries during the period of our analysis, and its members typically trade intensively with each other, which in turn makes the CU effect on the *Partner's PTA* variable to be large on its member countries.

5.1 Introducing Trade Costs into Our Model

While our model described in Section 2 already embodies tariffs (denoted by ts), it does not explicitly model other trade costs, such as transportation costs. In contrast to tariffs that generate revenues for governments, non-tariff trade costs, denote by τs, typically do not generate government revenues, thus enter into the welfare functions in a different way.[44] In particular, τ will affect firms' pricing (thus, their sales) in the same way that a tariff t affects it, without generating any tariff revenue. Because a country's decision to sign a new PTA (and to continue an existing one) would depend on the welfare change expected from signing it (or keeping it), once again we focus on this welfare change in this section, now denoted by $\triangle_j W^i_{\text{with } \tau}$ for country i considering to sign a new PTA (and to continue an existing one) with country j with non-tariff trade costs being included and denoted by a subscript, "with τ":

$$\begin{aligned} \triangle_j W^i_{\text{with } \tau} &= \triangle_j W^i \\ +\frac{s^i}{\mu^i}\mu^j\{\frac{-\tau^j_i}{2(1-\sigma)} &+ \frac{\sigma}{2(1-\sigma)(2-\sigma)}\overline{\tau}^j\}(1-\frac{\sigma s^i}{2-\sigma})t^j \\ -(1-\frac{3}{4(1-\sigma)})s^j\tau^i_j t^i &- \frac{\sigma\left(7-4\sigma+2s^i\right)}{4(1-\sigma)(2-\sigma)}\frac{s^j\sigma t^i}{2-\sigma}\overline{\tau}^i \end{aligned} \tag{21}$$

where $\triangle_j W^i$ is defined as in (10), τ^i_j represents the non-tariff trade costs of shipping products from country j to country i, and $\overline{\tau}^i$ denote the average non-tariff trade costs to country i with $\overline{\tau}^i \equiv \sum_{j=1}^{n} s^j \tau^i_j$

5.2 Construction of Variables based on Data

Prior to calibrating our model, we need to construct the variables of our model included in (21) based on the real data. The following Table 8 summarizes how we construct the variables.

As indicated in Table 8, data on tariffs and non-tariff trade costs are available in the form of ad valorem rates instead of specific ones. Thus, we need to covert data

[44] Some non-tariff trade barriers, such as anti-dumping or countervailing duties, may generate tariff revenues. But, such non-tariff barriers are contingent upon the conditions that justify such protection, which makes an introduction of such protection measures into our model beyond the scope of this paper.

[45] As shown later, we also use a country's merchandise export share of the world for s^i as a robustness check.

Table 8: Data description for calibration

Variable	Data source
s^i : GDP share of country i in the world[45]	WB Indicators
μ^i : Population of country i	WB Indicators
t^i_j: Specific tariff converted from ad valorem tariff	WTO
τ^i_j : Specific non-tariff cost converted from ad valorem non-tariff cost	ESCAP WB trade cost
$s^{\widehat{C}_i}$: GDP ratio of PTA partners of country i	WB Indicators

on valorem rates into specific ones in a way that is compatible with our model. Referring the full description of the numerical approximation method for such conversion to Appendix 2, we provide a brief discussion of the conversion method in this subsection.

To make the derived specific tariffs and non-tariff trade costs be suitable for the welfare change calculation associated with signing a new PTA in (21), we use the following equality that requires country i's consumer price of its import from country k to be the same regardless of the form of tariffs and non-tariff trade costs being utilized:

$$\left(1 + t^i_{ak} + \tau^i_{ak}\right) p^i_k\left(\cdot\right) = p^i_k\left(\cdot\right) + t^i_k + \tau^i_k, \tag{22}$$

where t^i_{ak} and τ^i_{ak} denote the ad valorem tariffs and non-tariff trade costs, respectively, and $p^i_k\left(\cdot\right)$ represents the exporting price of country k's firm that exports its products to country i. Even though $p^i_k\left(\cdot\right)$ is not directly observable, the profit maximization in our model with trade costs yields:

$$p^i_k\left(\cdot\right) = \frac{1-\sigma}{2-\sigma} - \frac{1}{2}t^i_k - \frac{1}{2}\tau^i_k + \frac{\sigma}{2(2-\sigma)}\left(\bar{t}^i + \bar{\tau}^i\right). \tag{23}$$

Using (22) and (23), one may derive $t^i_k + \tau^i_k$ as a function of $t^i_{ak} + \tau^i_{ak}$, $\bar{t}^i + \bar{\tau}^i$, and σ. Based on this function, Appendix 2 derives the numerical approximation method for finding the vector of approximated values of $t^i_k + \tau^i_k$ based on the vector of observed values of $t^i_{ak} + \tau^i_{ak}$. Because $t^i_k = t^i_{ak}p^i_k\left(\cdot\right)$ and $\tau^i_k = \tau^i_{ak}p^i_k\left(\cdot\right)$ from (22), then we can obtain the numerically approximated values for t^i_k and τ^i_k, separately, based on the numerically approximated value for $p^i_k\left(\cdot\right)$.

5.3 Calibration of Parameters

With regard to calibrating our model's parameter(s), we only have one parameter that is not determined yet in the baseline welfare change of country i from signing (or keeping) a PTA with country j of (21), σ. Because this parameter represents the consumer's willingness to substitute between different varieties, it is natural to use an estimate (or a range of estimates) of the parameter based on a regression analysis that utilizes the information on how countries' import volumes change in response to changes in their import prices. Even though there exists a sizable number of studies on estimating the substitutability of import demand based on the CES utility function, there is no such a empirical study on the import demand based on a quasi-linear utility function similar to the one in (1) to our knowledge.[46]

Instead of developing an empirical method to estimate the substitutability parameter, we run the following numerical excise on σ to check how a change in σ affects the countries' incentive to sign or keep PTAs in our model, thus its predictability of PTA formation between countries. This numerical exercise (and the calibration of extended models of ours below) are conducted through the following procedure. For a fixed value of σ that ranges from 0.1 to 0.9 in the step of 0.1, first, we convert the ad valorem trade cost data into specific tariff and non-tariff trade cost variables, as described in Section 5.2. Then, we calculate welfare changes for all possible pairs of countries from signing a new PTA or keeping a pre-existing one, defined as in (21) using the constructed variables of our model. Only when the calculated welfare change is positive for both countries of a pair of potential or current PTA partners, we record that our model predicts a new PTA or an on-going PTA for the pair, and otherwise, we record that our model predicts no PTA for the pair. For nine different values of σ, the following table shows the number (and percentage) of correctly predicted cases both for a PTA being signed or kept and for no PTA being signed or kept.

Out of 313,384 observations (i.e., country pairs) without missing values from our data of 183 countries for 25 years from 1993 to 2017, we have 263,300 pairs with no PTA and 50,084 pairs with a PTA. For $\sigma = 0.1$, Table 9 shows that our

[46] There does exist the industrial organization literature that explores the issue of estimating the substitutability parameter of a variety of quasi-linear utility functions using firm-level or individual-level demand data for differentiated products. Applicability of such estimates to our model is very limited because our substitutability parameter is supposed to reflect the import substitutability across very broadly defined differentiated goods (in fact, encompassing all manufactured imports), which affects how countries' import volumes respond to changes in import prices.

Table 9: PTA predictability of the base model

With 313,384 observations			
sigma	(real, predicted)=(1,1)	(real, predicted)=(0,0)	overall fitness
0.1	6,168 (12%)	261,448 (99%)	85.4%
0.2	9,224 (18%)	257,392 (98%)	85.1%
0.3	13,486 (27%)	246,714 (94%)	83.0%
0.4	19,208 (38%)	223,014 (85%)	77.3%
0.5	25,792 (51%)	185,590 (70%)	67.5%
0.6	32,650 (65%)	112,696 (43%)	46.4%
0.7	38,790 (77%)	63,492 (24%)	32.6%
0.8	41.420(83%)	45,360 (17%)	27.7%
0.9	41,800 (83%)	45,002 (17%)	27.7%

base model correctly predicts 12% of PTAs being signed or kept, and correctly predicts 99% of no PTA being signed or kept. This bias in predicting no PTA correctly changes monotonically as σ increases, having more PTAs be predicted correctly and fewer no PTA cases be predicted correctly. For $\sigma = 0.9$, the bias in our model's PTA prediction is just the opposite, with the prediction success rate increasing to 83% for PTA cases and the one for no PTA cases decreasing to 17%.

Therefore, the numerical analysis indicates that a higher value of σ mostly raises countries' incentive to sign or keep PTAs. A higher value of σ implies higher substitutability between differentiated products that a country may import from her trading partners, which in turn implies a bigger impact of a PTA on its member regardless of whether the PTA's net welfare effect on its member is positive or negative. Given the economic environments in the world during 1993-2017 implied by our constructed variables in Section 5.2, thus, higher substitutability would have mostly strengthened the positive effects of PTAs relative to the negative ones among potential and real PTA partners, inducing them to sign or keep PTAs. While Table 9 does provide a column of "overall fitness" with the highest value for the lowest substitutability, it simply reflects that the number of country pairs with no PTA is more than 5 times the number of pairs with PTAs in our observation.[47]

[47] The overall fitness is a weighted average of the predictability percentages for PTA and no PTA cases, with the weight being the proportion of PTA and no PTA cases in the total observations.

5.3.1 Heterogeneity in Countries' PTA Preferences

Now we extend our base model into the one that allows countries to have heterogeneous preferences toward signing or keeping PTAs, that are not reflected in (and independent from) the net welfare changes implied by the base model in (21). In particular, we introduce a country specific PTA preference parameter, cv_i for country i, so that country i and country j sign or keep a PTA if and only if $\triangle_j W^i_{\text{with }\tau} > cv_i$ and $\triangle_i W^j_{\text{with }\tau} > cv_j$. Thus, the higher the value of cv_i, the less willing country i is to sign or keep a PTA with any trading partner.

The basic procedure of calibrating this extended model is practically the same as the one employed for our numerical excise for calculating the PTA predictability of our model in Table 9. The only difference is to introduce an additional stage of changing the country specific PTA parameters to maximize the overall fitness of this extended model.[48] The following Table 10 shows the results from this calibration.

Table 10: Calibration with a Country Specific PTA Preference

With 313,384 observations			
sigma	(real, predicted)=(1,1)	(real, predicted)=(0,0)	overall fitness
0.1	29,536 (59%)	257,455 (98%)	91.6%
0.2	28,744 (57%)	258,210 (98%)	91.6%
0.3	29,686 (59%)	257,717 (98%)	91.7%
0.4	29,011 (58%)	258,149 (98%)	91.6%
0.5	26,288 (52%)	259,332 (98%)	91.1%
0.6	26,172 (52%)	259,472 (99%)	91.1%
0.7	26,150 (52%)	259,476 (99%)	91.1%
0.8	25,708 (51%)	259,323 (98%)	91.0%
0.9	26,489 (53%)	259,133 (98%)	91.1%

As one can easily predict, the overall fitness of this extended model improves over the base model's. One noticeable aspect of the calibration results in Table

[48]The overall fitness of the model is defined as in footnote 47. By replacing the critical value with updated one, the same process was repeated until the critical value for each country converges to a certain number (i.e., no longer changes).

10 is that the overall fitness is very similar across different values of σ, ranging between 91% and 91.7%. This result, which is distinct from the one from the base model, shows that our extended model's improved predictability largely comes from the variation in country specific PTA preferences rather than a change in the consumer's substitutability parameter. Even though our extended model's overall fitness reaches its maximum when $\sigma = 0.3$ with its predictability for PTAs being 59% and for no PTAs being 98%, there is no basis for claiming that $\sigma = 0.3$ is a correct parameter.

5.3.2 Heterogeneity in Country Pairs' PTA Preferences

As another calibration exercise, we consider the possibility that the PTA preference is country-pair specific in the following way: A PTA is predicted between country i and j if and only if $\triangle_j W^i_{\text{with } \tau} > cv_{ij}$ and $\triangle_i W^j_{\text{with } \tau} > cv_{ji}$, with $cv_{ij} \neq cv_{ji}$ being allowed. Once we introduce this country-pair specific PTA preference, we need to drop the observations of country pairs that experience no change in their PTA status during 1993 - 2017. This is because the calibrated PTA preference parameters of such pairs will take either arbitrary large or small numbers, achieving perfect PTA predictability for them. The number of observations with country pairs' switching their PTA status is 64,504 with 33,718 PTA cases and 30,786 no PTA cases.

The procedure of calibrating this model is practically the same as the calibration employed for Table 10, generating the following results in Table 11.

Table 11: With 64,504 observations

With PTA status switching data			
sigma	(real, predicted)=(1,1)	(real, predicted)=(0,0)	overall fitness
0.1	30,950 (92%)	27,606 (90%)	90.78%
0.2	30,990 (92%)	27,583 (90%)	90.81%
0.3	31,001 (92%)	27,550 (89%)	90.77%
0.4	30,857 (92%)	27,604 (90%)	90.6%
0.5	30,629 (91%)	27,354 (89%)	89.9%
0.6	29,668 (88%)	26,358 (86%)	86.9%
0.7	29,091 (86%)	25,343 (82%)	84.4%
0.8	28,898 (86%)	24,740 (80%)	83.2%
0.9	28,834 (86%)	24,573 (80%)	82.8%

Despite the fact that we are employing a much larger number of parameters

(33,306 = 183×182 versus 183) to maximize the overall fitness of this extended model for a smaller number of observations (64,504 versus 313,384), the resulting overall fitness in Table 10 is smaller than the one in Table 9 for each level of the substitutability parameter. This somewhat surprising result comes from the fact that the observations for calibration in Table 11 contain only country pairs whose PTA status changes at least once during 1993 - 2017 and one pair of PTA preference values of each pair of countries should be calibrated to explain the change(s) in the PTA status of that pair during this period.

The predictability of the model is rather balanced for PTA cases and no PTA cases, reaching over 90% for the value of σ that is less than half. Another noticeable difference between the results in Table 11 and those of Table 10 is that the model's overall fitness reaches its maximum for $\sigma = 0.2$, then it decreases significantly, especially for high values of σ. Once again, however, there is no basis for claiming that $\sigma = 0.2$ is a correct parameter or for claiming that a high value of σ is less likely to be correct.

As a robustness check of our calibration analysis, we conduct it by using country i's merchandise export share of the world for the construction of s^i, instead of using country i's GDP share of the world.[49] This changes the number of observations due to the different availability of merchandise export data from that of GDP data, from 313,384 to 141,256 for the analysis with country specific PTA preferences, and from 64,504 to 64,542 for the analysis with country-pair specific PTA preferences. Despite these changes, the predictability of our models roughly stay the same, as shown in Table 12 for $\sigma = 0.1$.

Table 12: A robustness check with s^i being country i's export share of the world

With country specific PTA preferences				
sigma	(real, predicted)=(1,1)	(real, predicted)=(0,0)	overall fitness	data type
0.1	29,536 (59%)	257,455 (98%)	91.6%	GDP
0.1	11,863 (54%)	117,141 (98%)	91.3%	Merchandise export
With country-pair specific PTA preferences				
sigma	(real, predicted)=(1,1)	(real, predicted)=(0,0)	overall fitness	data type
0.1	30,950 (92%)	27,606 (90%)	90.8%	GDP
0.1	30,321 (90%)	28,117 (91%)	90.5%	Merchandise export

[49] Recall that s^j is defined as the mass of differentiated-product-producing firms in country j (out of the total mass of such firms of the world that is normalized to one) in our theoretical model of Section 2.

6 Conclusion

Based on a n-country trade model of Furusawa and Konishi (2007), our study derives the testable empirical model that can distinguish the loss sharing effect (of own pre-existing PTAs) from the concession erosion effect (of partner's pre-existing PTAs) on the PTA formation, initially identified by Chen and Joshi (2010). This model enables us to include all pre-existing PTA effect variables of the trade literature on the PTA formation in our regression analysis, which in turn demonstrates the empirical existence of "snowballing alongside domino effects" in the recent proliferation of PTAs. Another finding is the negative *Partner's PTA* effect on the PTA formation, especially through CU members who are required to jointly negotiate new PTAs, which in turn magnifies the *Partner's PTA* effect. This result of CUs' having a negative effect on the proliferation of PTAs empirically confirms the prediction of a negative role that the unanimity requirement of CUs may play in expanding their PTA membership, offered by the theoretical studies of comparing CUs with FTAs on the PTA formation.

We also conduct some calibration exercises of fitting our model with the actual formation of PTAs. The calibration demonstrates that our basic model's prediction success rate is somewhat limited: 99% for no PTA cases and 12% for PTA cases when $\sigma = 0.1$, which monotonically changes as σ increases, having 17% for no PTA cases and 83% for PTA cases when $\sigma = 0.9$. Once we extend our model by allowing each country or each pair of countries to have an individual PTA preference or pair-specific PTA preferences that are different from each other's, respectively, then the calibration starts to generate overall predictive success rates that are higher than 90%.

One possible future research direction is to extend our theoretical model to analyze a political economy aspect in the determination of trade policies, such as the one that the domino theory emphasizes, in addition to the *Own* and *Partner's PTA* effects that come from the pure welfare consideration. Such an extension will enable us to construct the pre-existing PTA variables based on an unified theoretical model, which in turn may sharpen the results of corresponding empirical analysis. This will also enable us to conduct the corresponding calibration exercises quantifying the relative importance of different pre-existing PTA effects in the proliferation of PTAs.

Appendix 1

This appendix explains why we choose to focus on the terms outside the curly brackets in (14) in constructing the pre-existing PTA effect variables as in (15) and (16). First, the common first term inside the curly brackets in (14), $\sigma/2(1-\sigma)(2-\sigma)$, measures how the substitutability parameter, σ, affects these pre-existing PTA effects: A higher σ raises these effects as a higher substitutability between imports from different countries strengthens the market access effect of having a PTA. Note that it affects all pre-existing PTA effects in the same magnitude, regardless of which pair of countries being considered for signing a PTA.[50] This justifies our ignoring the term in constructing the pre-existing PTA effect variables for our empirical analysis.

The common second term inside the curly brackets in (14), $(t^i)^2$, also magnifies the pre-existing PTA effects as the substitutability does: A higher MFN tariff of country i raises country i's pre-existing PTA effects. Differently from the common first term (i.e., $\sigma/2(1-\sigma)(2-\sigma)$), $(t^i)^2$ does vary across countries, possibly necessitating the inclusion of country i's tariff-level variable in constructing the pre-existing PTA effect variables. However, note that there exists a simultaneity problem between a country's pre-existing PTAs and her MFN tariff that is similar to the one between pre-existing FTAs and import shares: More pre-existing PTAs of a country may raise or reduce her MFN tariff on her non-PTA members as the previous studies on this issue demonstrate.[51] In contrast to the import share variables, thus the use of estimated and fixed MFN tariff variables can be contentious. In addition to having this hard-to-deal-with simultaneity issue, a country's tariff level can be strongly correlated with her preference toward a protective trade policy, affecting her incentive to sign a PTA.[52] Thus, we decide not to include them in constructing our pre-existing PTA effect variables.

With regard to the last terms in the curly brackets in (14), we cannot come up with any intuitive explanations for them. In addition, the measure of differentiated-product-producing firms that are located in country i or that in country j are multiplied with some ratio functions of the substitutability parameter, which makes constructing corresponding variables very hard. One bright side

[50] As shown in Section 5 of calibrating our model, a higher value for σ does raise the prediction success rate for a PTA being signed or kept and lower the prediction success rate for a PTA being not signed (presumably for any pair of countries in our data) in the our base model analysis.

[51] For example, See Limão (2016)for a detailed discussion of these studies.

[52] As one of our robustness checks, we include the bound (mostly negotiated through the Uruguay round negotiation and fixed) tariff variable in the regression, finding a statistically significant negative coefficient estimate on it.

is that these last terms in the curly brackets do not systematically undermine the validity of the constructed variables in (15) and (16): The last term for the Own PTA effect in (14) is increasing in s^i, thus not directly affecting $s^j s^{\widehat{C}_i}$, on which our Own PTA effect variable is constructed; a higher s^j decreases the whole Partner's PTA effect term in (14) for $s^j < 0.5$ (supposedly valid for any country in our data) as it does $-\left(\mu^i/\mu^j\right) s^j s^{\widehat{C}_i}$, on which our Partner's PTA effect variable is constructed.[53] Thus, we ignore these last terms within curly brackets in constructing our pre-existing PTA effect variables.

Appendix 2

We derive the numerical procedure to convert ad valorem tariffs and non-tariff trade costs associated with shipping products from country k to country i, denoted respectively by t^i_{ak} and τ^i_{ak}, into corresponding specific tariffs and non-tariff trade costs, t^i_k and τ^i_k in a model consistent way. Country i's consumer price of its import from country k should be the same regardless of forms of trade costs to make the welfare calculation be identical across different forms of trade costs, as discussed in Section 5.2, thus, the following equalities need to be satisfied:

$$\left(1 + t^i_{ak} + \tau^i_{ak}\right) p^i_k\left(\cdot\right) = p^i_k\left(\cdot\right) + t^i_k + \tau^i_k = \tfrac{1-\sigma}{2-\sigma} + \tfrac{1}{2} t^i_k + \tfrac{1}{2}\tau^i_k + \tfrac{\sigma}{2(2-\sigma)}\left(\bar{t}^i + \bar{\tau}^i\right),$$

where $p^i_k\left(\cdot\right)$ represents the exporting price of country k's firm that exports its products to country i. The latter equality comes from the profit-maximizing $p^i_k\left(\cdot\right)$ in our model, which is determined by the associated trade costs, t^i_k and τ^i_k, together with country i's market competition level that is influenced by $\bar{t}^i + \bar{\tau}^i$. The above equalities imply:

$$t^i_k + \tau^i_k = \left(t^i_{ak} + \tau^i_{ak}\right) p^i_k\left(\cdot\right) = \left(t^i_{ak} + \tau^i_{ak}\right)\left[\tfrac{1-\sigma}{2-\sigma} - \tfrac{1}{2}\left(t^i_k + \tau^i_k\right) + \tfrac{\sigma}{2(2-\sigma)}\sum_{j=1}^{n} s^j\left(t^i_j + \tau^i_j\right)\right]. \tag{24}$$

To convert $t^i_{ak} + \tau^i_{ak}$ into $t^i_k + \tau^i_k$, we use a numerical approximation method derived as follows. For i, we start with $k = 1\ (\neq i)$ (for $i = 1$,we start with $k = 2$),

[53] The derivative of Partner's PTA effect in (14) with respect to s^j is equal to $-\frac{\mu^i}{\mu^j} s^{\widehat{C}_i} \frac{\sigma}{2(1-\sigma)(2-\sigma)^2} \left(t^i\right)^2 \left[1 - \left(0.5 - s^j\right)\sigma\right]$.

converting the above formula into the following one, using the assumption of $t_k^i + \tau_k^i = T_0^i \in (0,1)$ for all k ($\neq i$; note that $t_i^i + \tau_i^i = 0$) with T_0^i denoting an arbitrary value for an initial value for numerical approximation, such as 0.5, as country i's average trade cost:

$$\left(t_1^i + \tau_1^i\right)_0 = \frac{\left[2\left(1-\sigma\right) + \sigma\left(1 - s^1 - s^i\right) T_0^i\right]\left(t_{a1}^i + \tau_{a1}^i\right)}{2\left(2-\sigma\right) + \left[2 - \left(1+s^1\right)\sigma\right]\left(t_{a1}^i + \tau_{a1}^i\right)}, \tag{25}$$

with (25) being derived from

$$\left(t_1^i + \tau_1^i\right)_0 = \left(t_{a1}^i + \tau_{a1}^i\right)\left[\tfrac{1-\sigma}{2-\sigma} - \tfrac{1}{2}\left(t_1^i + \tau_1^i\right)_0 + \tfrac{\sigma s^1}{2(2-\sigma)}\left(t_1^i + \tau_1^i\right)_0 + \tfrac{\sigma}{2(2-\sigma)}\sum_{j=2(\neq i)}^{n} s^j T_0^i\right].$$

It is easy to check that that $\partial\left(t_1^i + \tau_1^i\right)_0 / \partial\left(t_{a1}^i + \tau_{a1}^i\right) > 0$ and $\left(t_1^i + \tau_1^i\right)_0 \in (0,1)$.

Then, the next step is to calculate an initial numerically approximated value for $t_2^i + \tau_2^i$, denoted by $\left(t_2^i + \tau_2^i\right)_0$ as follows:

$$\left(t_2^i + \tau_2^i\right)_0 = \frac{\left[2\left(1-\sigma\right) + \sigma s^1\left(t_1^i + \tau_1^i\right)_0 + \sigma\left(1 - s^1 - s^2 - s^i\right) T_0^i\right]\left(t_{a2}^i + \tau_{a2}^i\right)}{2\left(2-\sigma\right) + \left[2 - \left(1+s^2\right)\sigma\right]\left(t_{a2}^i + \tau_{a2}^i\right)}, \tag{26}$$

with (26) being derived from

$$\begin{aligned}\left(t_2^i + \tau_2^i\right)_0 = \left(t_{a2}^i + \tau_{a2}^i\right)\Big[&\tfrac{1-\sigma}{2-\sigma} - \tfrac{1}{2}\left(t_2^i + \tau_2^i\right)_0 + \tfrac{\sigma s^2}{2(2-\sigma)}\left(t_2^i + \tau_2^i\right)_0 + \\ &\tfrac{\sigma s^1}{2(2-\sigma)}\left(t_1^i + \tau_1^i\right)_0 + \tfrac{\sigma}{2(2-\sigma)}\sum_{j=3(\neq i)}^{n} s^j T_0^i\Big].\end{aligned}$$

In a similar manner we can calculate an initial numerically approximated value for $t_{sk}^i + t_{rk}^i$, denoted by $\left(t_{sk}^i + t_{rk}^i\right)_0$ as follows:

$$\left(t_k^i + \tau_k^i\right)_0 = \frac{\left[2\left(1-\sigma\right) + \sigma\sum_{j=1(\neq i)}^{k-1} s^j\left(t_j^i + \tau_j^i\right)_0 + \sigma\left(1 - \sum_{j=1(\neq i)}^{k} s^j - s^i\right) T_0^i\right]\left(t_{ak}^i + \tau_{ak}^i\right)}{2\left(2-\sigma\right) + \left[2 - \left(1+s^k\right)\sigma\right]\left(t_{ak}^i + \tau_{ak}^i\right)}. \tag{27}$$

Once again, we can easily show that $\partial\left(t^i_{sk}+t^i_{rk}\right)_0/\partial\left(\tau^i_{ak}+\tau^i_{rk}\right) > 0$ and $\left(t^i_{sk}+t^i_{rk}\right)_0 \in (0,1)$, with

$$\sigma\left(1-s^k-\sum_{j=1(\neq i)}^{k-1} s^j - s^i\right)+\sigma\left(\sum_{j=1(\neq i)}^{k-1} s^j + s^i\right)-$$
$$\sigma\left(1-\sum_{j=1(\neq i)}^{k} s^j - s^i\right)T^i_0-\sigma\sum_{j=1(\neq i)}^{k-1} s^j\left(t^i_{sj}+t^i_{rj}\right)_0 > 0.$$

Using the above formula in (27), we can calculate the initial numerically approximated value for country i's total trade cost with country k, $\left(t^i_k+\tau^i_k\right)_0$ for all $k \neq i$. Then, we define $T^i_1 \equiv \sum_{j=1(\neq i)}^{n} s^j\left(t^i_j+\tau^i_j\right)_0$, and calculate the "first" numerically approximated values for country i's trade cost with country k, $\left(t^i_{sk}+t^i_{rk}\right)_1$, using the following formula:

$$\left(t^i_k+\tau^i_k\right)_1=\frac{\left[2(1-\sigma)+\sigma\sum_{j=1(\neq i)}^{k-1} s^j\left(t^i_j+\tau^i_j\right)_1+\sigma\left(1-\sum_{j=1(\neq i)}^{k} s^j-s^i\right)T^i_1\right]\left(t^i_{ak}+\tau^i_{ak}\right)}{2(2-\sigma)+\left[2-\left(1+s^k\right)\sigma\right]\left(t^i_{ak}+\tau^i_{ak}\right)}, \tag{28}$$

once again, starting from $k=1\,(\neq j)$ to $k=n\,(\neq j)$. Then, we can calculate the following measure of approximation level, denoted by $AL\,(1)$

$$AL\,(1)=\sum_{j=1(\neq i)}^{n(\neq i)}\left|\left(t^i_j+\tau^i_j\right)_1-\left(t^i_j+\tau^i_j\right)_0\right|. \tag{29}$$

If $AL\,(1) < c$, a critical value for ending the numerical approximation, let's say 0.0001, then we define our numerically approximated values of country i's trade costs by $t^i_k+\tau^i_k \equiv \left(t^i_k+\tau^i_k\right)_1$ for all $k \neq i$.

If $AD\,(1) \geq c$, then we define $T^i_2 \equiv \sum_{j=1(\neq i)}^{n} s^j\left(t^i_j+\tau^i_j\right)_1$, and calculate the "second" numerically approximated values for country i's trade cost with country k, $\left(t^i_{sk}+t^i_{rk}\right)_2$, in the same manner as in (28) with T^i_2 replacing T^i_1. Then, we can calculate the following measure of approximation degree, denoted by $AD\,(2)$,

$$AD\,(2)=\sum_{j=1(\neq i)}^{n(\neq i)}\left|\left(t^i_{sj}+t^i_{rj}\right)_2-\left(t^i_{sj}+t^i_{rj}\right)_1\right|.$$

If $AD(2) < c$, a critical value for ending the numerical approximation, let's say 0.0001, then we define our numerically approximated values of country i's trade costs by $t^i_k + \tau^i_k \equiv \left(t^i_k + \tau^i_k\right)_2$ for all $k \neq i$. If $AD(2) \geq c$, then we repeat the above approximation procedure until we have $AD(m) < c$. Once we find such a condition being met with m-th iteration, then we define our numerically approximated values of country i's trade costs by $t^i_k + \tau^i_k \equiv \left(t^i_k + \tau^i_k\right)_m$ all $k \neq i$.

Now, we derive the numerical approximation formula for country i's MFN specific tariff as follows. From (24), $t^i_k = t^i_{ak} p^i_k(\cdot)$ and $\tau^i_k = \tau^i_{ak} p^i_k(\cdot)$. Recall that we have defined our numerically approximated values of country i's total specific trade costs by $t^i_k + \tau^i_k \equiv \left(t^i_k + \tau^i_k\right)_m$ all $k \neq i$, which in turn implies that we can obtain our numerically approximated value of country k's firm's export price to country i, denoted by $\left[p^i_k(\cdot)\right]_m$, as follows:

$$\left[p^i_k(\cdot)\right]_m = \frac{1-\sigma}{2-\sigma} - \frac{2-\left(1+s^k\right)\sigma}{2(2-\sigma)}\left(t^i_k + \tau^i_k\right)_m + \frac{\sigma}{2(2-\sigma)} \sum_{j=1(\neq i)}^{k-1} s^j \left(t^i_j + \tau^i_j\right)_m + \frac{\sigma\left(1 - \sum_{j=1(\neq i)}^{k} s^j - s^i\right)}{2(2-\sigma)} \left(T^i_m\right). \tag{30}$$

Then, we can define our numerically approximated value of country i's specific tariff on its import from country k by $\left(t^i_k\right)_m = t^i_{ak}\left[p^i_k(\cdot)\right]_m$. Finally, we obtain our numerically approximated value of country i's specific non-tariff trade costs on its import from country k from $\left(\tau^i_k\right)_m = \left(t^i_k + \tau^i_k\right)_m - \left(t^i_k\right)_m$.

There are two things to discuss about the above approximation. First, note that the numerically approximated value of country k's firm's export price to country i in (30) utilizes the same procedural numerical value of $p^i_k(\cdot)$ that we use for our calculation of $\left(t^i_k + \tau^i_k\right)_m$. Second, it is possible to have $t^i_k \neq t^i_j$ even when $t^i_{ak} = t^i_{aj}$ for $k \neq j$. This is because $p^i_k(\cdot) \neq p^i_j(\cdot)$ unless $t^i_k + \tau^i_k = t^i_j + \tau^i_j$.

References

Baier, S. L. and J. H. Bergstrand (2004). Economic determinants of free trade agreements. *Journal of international Economics 64*(1), 29–63.

Baier, S. L., J. H. Bergstrand, and R. Mariutto (2014). Economic determinants of free trade agreements revisited: Distinguishing sources of interdependence. *Review of International Economics 22*(1), 31–58.

Baldwin, R. (1993). A domino theory of regionalism.

Baldwin, R. and D. Jaimovich (2012). Are free trade agreements contagious? *Journal of international Economics 88*(1), 1–16.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies 47*(1), 225–238.

Chen, M. X. and S. Joshi (2010). Third-country effects on the formation of free trade agreements. *Journal of International Economics 82*(2), 238–248.

Corden, W. (1974). *Trade Policy and Economic Welfare*. Clarendon Press.

Egger, P. and M. Larch (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics 76*(2), 384–399.

Furusawa, T. and H. Konishi (2007). Free trade networks. *Journal of International Economics 72*(2), 310–335.

Jackson, M. O. and A. Wolinsky (1996). A strategic model of social and economic networks. *Journal of economic theory 71*(1), 44–74.

Krugman, P., A. J. Venables, et al. (1990). Integration and the competitiveness of peripheral industry. Technical report, CEPR Discussion Papers.

Limão, N. (2016). Preferential trade agreements. In *Handbook of commercial policy*, Volume 1, pp. 279–367. Elsevier.

Neumayer, E. and T. Plümper (2010). Spatial effects in dyadic data. *International Organization 64*(1), 145–166.

Ottaviano, G., T. Tabuchi, and J.-F. Thisse (2002). Agglomeration and trade revisited. *International economic review*, 409–435.

Ovádek, M. and I. Willemyns (2019). International law of customs unions: Conceptual variety, legal ambiguity and diverse practice. *European Journal of International Law 30*(2), 361–389.

Saggi, K., A. Woodland, and H. M. Yildiz (2013). On the relationship between preferential and multilateral trade liberalization: the case of customs unions. *American Economic Journal: Microeconomics 5*(1), 63–99.

Saggi, K. and H. M. Yildiz (2010). Bilateralism, multilateralism, and the quest for global free trade. In *Economic Analysis of the Rules and Regulations of the World Trade Organization*, pp. 156–167. World Scientific.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data mit press. *Cambridge, MA 108*.

Yi, S.-S. (1996). Endogenous formation of customs unions under imperfect competition: open regionalism is good. *Journal of International Economics 41*(1-2), 153–177.