# An Oligopolistic Theory of Regional Trade Agreements

Christian Soegaard*

20th August 2012

**Abstract**

This paper sets up a three-country model of oligopoly to analyse the relationship between trade costs and trade policy cooperation. Acting non-cooperatively, the three countries are caught in a prisoner's dilemma in which import tariffs are used to improve one country's terms of trade and to shift profits towards its domestic market at the expense of trading partners. Globally, the efficient trade policy depends critically on the nature of the gains from trade. If there is some degree of substitutability between goods, there is an inherent trade-off between pro-competitive gains from trade and losses in transit. If goods are independent, on the other hand, there is no such trade-off. The degree of substitutability, thus, becomes a crucial parameter in terms of determining the sustainability of regional versus global trade integration. Provided there is some degree of substitutability between goods, I find that there is scope for two countries to form a regional trade agreement, leaving the third country out whenever extra-regional trade costs exceed a critical threshold. If goods are independent of each other, however, there is no trade-off between pro-competition and losses in transit and trade costs become irrelevant to global integration. I also find that customs unions may be less regional than free trade areas due to their higher external tariffs.

**Keywords:** trade policy, self-enforceability, trade costs, free trade agreement.
**JEL codes:** F13, F15.

*The Leverhulme Centre for Research on Globalisation and Economic Policy, University of Nottingham,UK. Email: lexcs4@nottingham.ac.uk.

# 1 Introduction

The history of trade liberalisation in the post-war era is intimately related with the expansion of the GATT/WTO, and to the signing of a countless number of bilateral and regional trade agreements. Since World War II, average ad valorem import tariffs have been reduced from over 40 percent to less than 4 percent, and according to Freund and Ornelas (2010) the average WTO member has now signed regional or bilateral trade agreements with fifteen countries. There are clearly strong forces pushing countries to sign trade agreements, and it is important for economists and political scientists alike to understand the nature and causes of the desire to engage in cooperative trade policy. Why do countries sign trade agreements, and what determines the extent of trade liberalisation?

While the trade policy literature offers many explanations for why trade agreements are preferential, an interesting yet under-researched question relates to why preferential trade agreements (PTA) are regional. The present paper contributes to this research question by examining the relationship between natural barriers to trade, such as natural distance and transport costs, and those that are determined at the political level such as import tariffs.

To the best of my knowledge, Krugman (1991) is the first to address the regional bias of PTAs. To illustrate his point, he uses an extreme example: suppose trade costs are zero between countries on the same continent and prohibitive between countries on different continents. In this case, an inter-continental trade agreement would offer no benefits at all, while a PTA would liberalise all trade which is possible to liberalise and thus raise welfare. In this example the continent can be considered a natural trading block.

As Ludema (2002) notes, however, in the discussion surrounding natural trade blocks, it may not be assumed that governments actually enter into the trade agreements that yield the highest welfare. He constructs a model based on repeated games which offers an explanation for the observed regional bias. In his framework trade policy cooperation is limited by a constraint that balances the one-time incentive to deviate from a trade agreement with the discounted benefits of future cooperation. In particular, he constructs a model of horizontal FDI in which firms face a trade-off between being close to their foreign markets and to concentrate production and exploit economies of scale. When trade costs are higher, firms have a greater incentive to switch production abroad. The motive behind imposing political trade restrictions in the form of import tariffs is increasing in trade costs, because with higher trade costs there is a higher unilateral incentive to attract firms to locate in the domestic country. This motive is commonly referred to as the home market effect in the trade policy literature. Ludema (2002) exploits the positive relationship between natural and political trade costs to generate a regional bias in the formation of free trade agreements. He shows that the benefits of cooperation rise, and the incentive to deviate falls as natural trade costs are reduced.

In the present paper I construct an infinitely repeated game in an oligopolistic model of trade with product diversification as in Yi (1996). The desire to impose unilateral trade restrictions is driven by a standard terms of trade motive in addition to a profit-shifting motive. Since entry and exit is ruled out, there is no home market effect, in the sense that it is not possible to attract firms to relocate to the home country. The present paper does not seek to rebut the findings in Ludema (2002). On the contrary, it seeks to examine whether his findings are robust to alternative rationales for trade policy. The fact that the present model features terms-of-trade and profit-shifting motives for trade policy also offers some advantages in a repeated game framework where enforcement is ensured through trigger strategies. This is because I would argue that the benefits of improved terms of trade and greater domestic profits when a country defects from an agreement could be expected to be more immediate (cite Feenstra's evidence on pass-through of exchange rates). In Ludema's (2002) model where benefits of unilateral trade policy are realised only when firms have physically relocated to the home country, it could be expected that such gains take so long that trading partners will have plenty of time to impose the appropriate punishment.

In the present model, the degree of product differentiation is a crucial parameter in terms of generating a regional bias in trade agreements. If goods are substitutable to some extent, there is an inherent trade-off between pro-competitive gains from trade and losses in transit due to natural, as opposed to political, trade costs. Hence, the welfare of a trade agreement is higher when trade costs are lower. In an infinitely repeated game, I demonstrate that a trade agreement that yields higher welfare is also more self-enforcing. A trade agreement is required to be self-enforceable whereby countries weigh their one-shot benefit from defection with the long-run cost of infinite Nash reversion. When natural trade costs are close to their prohibitive levels, the welfare loss that results from losses in cross-hauling outweigh the pro-competitive gains from trade and the costs of defecting from a trade agreement therefore become smaller relative to the gain from defection. Hence, the trade agreement is not self-enforcing for a high range of trade costs. When goods are independent of each other there is no pro-competitive gain from trade and the gain from trade is driven only by variety. In this case there is no trade-off between pro-competition and losses in transit and thus natural trade costs become irrelevant in determining the self-enforceability of trade agreements.

I construct the model first in a two-country framework whereby the link between natural trade costs and trade policy cooperation is established without the more complicated trade diverting effects that come about from a third country. In this two-country framework I find that free trade is more self-enforcing when trade costs fall provided there is a minimum degree of substitution between goods. I show that there are two thresholds of natural trade barriers which must be passed before free trade is sustainable and desirable. When trade costs are close to their prohibitive levels, free trade cannot be supported for any levels of the governments' discount factors. As trade costs fall there will be a range of trade costs in which free trade is enforceable but undesirable from a joint welfare-maximising perspective. However, as trade costs fall further and pass the second threshold, free trade is both desirable and enforceable as the pro-competitive gains from trade dominate the losses from cross-hauling. The relationship between natural trade costs and trade policy cooperation, however, is less pronounced when goods become more differentiated, and when goods are independent of each other, trade costs become irrelevant for the self-enforceability of free trade.

I extend the model to a three-country setting in which trade-diverting effects of regional trade agreements can also be considered. I assume there are two countries that are located close to each other with zero trade costs between them, and a third country which is further away but equi-distant from its two trading partners. In this case the two close trading partners may sign a PTA with one another or a global free trade agreement. I show that the patience required to sustain global free trade is greater for the distant country such that this country is pivotal in the determination of global trade policy cooperation. I compare the two cases of a preferential free trade agreement (FTA) and a preferential Customs Union (CU).

For the case of an FTA and where the degree of substitutability between goods is greater than zero I show that when trade costs fall between the distant partner and the other two, global free trade can be supported for a larger range of discount factors. The relative self-enforceability of global free trade and a preferential FTA, however, depends on whether deviation from global free trade implies contagion whereby every country in the world deviate from each other following the deviation by the pivotal country. If there is contagion the pivotal distant country is punished by the standard Nash tariff which is the tariff which maximises the other two countries' individual welfare. I show that in this case, for low trade costs, free trade can be supported for a larger range of discount factors, whereas when trade costs exceed a critical threshold there is scope for regionalism in the sense that a regional FTA can be supported for a larger range of discount factors. When there is no contagion such that the two close trading partners would continue as a regional FTA, however, the punishment experienced by the pivotal country is lower. This is due to the well-known tariff complementarity effect first identified by Richardson (1990 - check this was the first!). Countries in the regional FTA set lower external tariffs since there PTA works so as to divert trade away from the non-member. They compensate by setting lower external tariffs. In this case, it can be readily shown that the discount factor required to sustain global free trade is always higher than that required to sustain a preferential FTA. Hence, in the no contagion case there is scope for regionalism for every level of natural trade restrictions. When goods are independent of each other, there is no difference between the self-enforceability of either type of trade agreement, which implies that natural trade barriers are irrelevant to the formation of trade agreements.

For the CU case, and when there is a positive degree of substitution between goods I arrive at different results. While it is still the case that the self-enforceability of global free trade increases when trade costs fall, the pivotal distant country experiences a much harder punishment if there is no contagion. This is because the common external tariff of a CU exploits its increased monopoly power by setting higher external tariffs than is the case for an FTA. Thus, the critical discount factor required to sustain global free trade is much lower in the no contagion case. Hence, it can be shown that a CU has a potential to be far less regional than an FTA if there is no contagion.

For both an FTA and a CU, however, I find that a reassuring result which is that any trade agreement which is self-enforceable is also desirable.

The paper is organised as follows. In Section 2, I present the oligopolistic model with product differentiation which will be used throughout the paper. In Section 3, I present some trade policy preliminaries in addition to defining the repeated game. Section 4 presents the simple two-country case before Section 5 analyses the full triangular model. Finally, Section 6 offers some concluding remarks as well as some ideas for future research.

## 2 The model

In this section I present a version of Yi's (1996) extension of the Brander (1981) model of trade with oligopoly. I choose a simple setup with a world of $n \geq 2$ countries with one firm in each. There are an infinite number of discrete time periods, and each firm produces one good. Preferences are identical across countries and can be represented by the following quasilinear-quadratic utility function in each period:

$$U\left(\mathbf{q_i}, q_{i0}\right) = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2}\sum\nolimits_{j=1}^{n} q_{ij}^2 + q_{i0}, \quad i = 1..n, \tag{1}$$

where $q_{ij}$ is country $i$'s consumption of country $j$'s products, $\mathbf{q_i} \equiv (q_{i1}, q_{i2}, ..., q_{in})$ is country $i$'s consumption vector, $Q_i \equiv \sum_{j=1}^{n} q_{ij}$ and $q_{i0}$ is country $i$'s consumption of the numeraire good. The numeraire is freely traded across countries to settle the balance of trade. Labour is the only factor of production in this model and its marginal product is constant and normalised to unity. The labour endowment in each country is the same, and I assume that these endowments are large enough to guarantee a positive consumption of the numeraire good in equilibrium. These assumptions ensure that the wage rate is equal to one in equilibrium. The parameter $\gamma \in [0; 1]$ represents a substitution index: when $\gamma = 0$ goods are independent and each firm is a monopolist in its own market. As $\gamma$ increases goods become closer substitutes. Assuming $\gamma < 1$ consumers have a taste for variety. Notice that $\gamma$ can be thought of as a measure of the degree of strategic interaction between firms, such that a higher $\gamma$ implies a more direct competition among firms. Country $i$'s inverse demand for country $j$'s good can be derived by maximising utility in (1):

$$p_{ij} = a - (1-\gamma)q_{ij} - \gamma Q_i = a - q_{ij} - \gamma \sum_{\substack{k=1 \\ k \neq j}}^{n} q_{ij}. \tag{2}$$

Trade is subject to natural trade costs of the iceberg form. In order for one unit of exports to arrive in country $i$, $1 + \alpha_{ij}$ units must be produced. I assume there are no internal natural trade costs, $\alpha_{ii} = 0$, and that trade costs between any country pair $i$ and $j$ are symmetric, such that $\alpha_{ij} = \alpha_{ji}$. In addition to natural trade costs, the governments of each country are able to impose political trade costs in the form of a specific import tariff. I assume that tariffs are country-specific such that country $i$ sets a tariff equal to $\tau_{ij}$ on imports from country $j$. I also assume there are no internal political trade barriers, $\tau_{ii} = 0$.

All firms produce at the same marginal cost in terms of units of labour, $w$, in their respective domestic markets, but due to trade costs (both political and natural) the effective marginal cost of exporting to country $i$ becomes $w_{ij} = w + \alpha_{ij} + \tau_{ij}$ for the firm in country $j$. Markets are segmented and firms compete in a Cournot fashion by choosing quantities in each country. In country $i$ firm $j$ solves the following problem, $\max_{q_{ij}} \pi_{ij} = \left(p_{ij} - w_{ij}\right) q_{ij}$. This yields the first-order conditions:

$$p_{ij} - w_{ij} - q_{ij} = 0 \tag{3}$$

Using (2) these conditions can be rewritten as:

$$a - w - (2-\gamma)q_{ij} - \gamma Q_i = 0 \tag{4}$$

Summing the first-order conditions in (4) gives the following per-period quantities in Cournot equilibrium:

$$q_{ij} = \frac{\Gamma(0,\gamma) + \gamma(T_i + A_i) - \Gamma(n,\gamma)(\tau_{ij} + \alpha_{ij})}{\Gamma\left(0,\gamma\right)\Gamma\left(n,\gamma\right)}; \qquad Q_i = \frac{n - T_i - A_i}{\Gamma(n,\gamma)}, \tag{5}$$

where $T_i \equiv \sum_{j=1}^{n} \tau_{ij}$, $A_i \equiv \sum_{j=1}^{n} \alpha_{ij}$, and I have normalised such that $a - w = 1$. The function $\Gamma\left(\cdot\right)$ is defined as:

$$\Gamma\left(k,\gamma\right) \equiv 2 - \gamma + k\gamma. \tag{6}$$

The equilibrium quantities have the following properties:

$$\begin{aligned} \frac{dq_{ij}}{d\tau_{ij}} &= \frac{dq_{ij}}{d\alpha_{ij}} = \frac{\gamma - \Gamma(n,\gamma)}{\Gamma(0,\gamma)\Gamma(n,\gamma)} < 0; \quad \frac{dq_{ih}}{d\tau_{ij}} = \frac{dq_{ih}}{d\alpha_{ij}} = \frac{\gamma}{\Gamma(0,\gamma)\Gamma(n,\gamma)} > 0 \quad \text{for} \quad h \neq j; \\ \frac{dQ_i}{d\tau_{ij}} &= \frac{dQ_i}{d\alpha_{ij}} = -\frac{1}{\Gamma\left(n,\gamma\right)} < 0. \end{aligned} \tag{7}$$

If country $i$ raises its tariff on imports from country $j$, the consumption of good $j$ falls, but the consumption of all other goods increases. Total consumption, however, falls. Exogenous increases in natural trade costs, $\alpha_{ij}$, have the same effect on quantities. In fact, what matters for the equilibrium quantities are *total* trade costs whether political or natural.

Using the first-order condition in (3) I obtain an expression for the per-period equilibrium profits of firm $j$ in country $i$:

$$\pi_{ij} = (p_{ij} - w_{ij})q_{ij} = q_{ij}^2.$$

The equilibrium profits have the following properties:

$$\frac{d\pi_{ij}}{d\tau_{ij}} = \frac{d\pi_{ij}}{d\alpha_{ij}} = \frac{2(\gamma - \Gamma(n,\gamma))q_{ij}}{\Gamma(0,\gamma)\Gamma(n,\gamma)} < 0; \quad \frac{d\pi_{ih}}{d\tau_{ij}} = \frac{d\pi_{ih}}{d\alpha_{ij}} = \frac{2\gamma q_{ih}}{\Gamma(0,\gamma)\Gamma(n,\gamma)} > 0 \quad \text{for} \quad h \neq j.$$

If country $i$ raises its tariff (or there is an exogenous rise in natural trade costs) on imports from country $j$, the export profits of the firm $j$ fall, but country $i$'s firm's profits increase along with the profits of all other firms exporting to country $i$.

There are two sources of gains from trade in the model: an increased variety of goods and decreased market power of the domestic industry. When the substitution index $\gamma$ is lower, consumers value variety whereas the pro-competitive effect is higher when $\gamma$ is higher. Whether the gains from trade are of the pro-competitive type or of the variety type have important implications for the arguments that follow in this paper.

The welfare of each country consists of consumer surplus, tariff revenue, the sum of domestic and export profits and labour income. The expression for per-period consumer surplus takes the following form:

$$CS_i = \sum_{j=1}^{n} \frac{1}{2}(a - p_{ij})q_{ij}. \tag{8}$$

The per-period tariff revenue which is redistributed back to individuals in a lump-sum fashion is given as:

$$TR_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} \tau_{ij} q_{ij}. \tag{9}$$

The sum of domestic and export profits of firm $i$ is:

$$\Pi_i = \pi_{ii} + \sum_{\substack{j=1 \\ j\neq i}}^{n} \pi_{ji} = (p_{ii} - w)q_{ii} + \sum_{\substack{j=1 \\ j\neq i}}^{n} (p_{ji} - w - \tau_{ji} - \alpha_{ji})q_{ji}, \tag{10}$$

and finally, labour income is:

$$\ell_i = wL_i, \tag{11}$$

where $L$ is the total number of workers. The per-period welfare of country $i$ can thus be expressed by adding up (8)-(11):

$$W_i \equiv CS_i + TR_i + \Pi_i + \ell_i \tag{12}$$

# 3 Trade policy

In this section I present some preliminaries of trade policy which will serve as a point of departure for the discussion that follows. What drives countries to impose import tariffs unilaterally, and why do they wish to cooperate? To answer these questions, I follow the literature on trade policy (see for example Bagwell and Staiger (2010) for an overview), and discuss how trade costs affect the objectives.

### 3.1 Unilateral trade policy

When acting non-cooperatively, it is assumed that the governments of each country set tariffs so as to maximise their individual welfare. The governments move first by setting optimal tariffs and firms then set Cournot quantities subject to the tariffs chosen by the governments in each market. As discussed in Baldwin and Venables (1995) and Mrazova (2009), it is possible to decompose the welfare effects of import tariffs into a terms-of-trade effect (ToT), a volume-of-trade effect (VoT), and a profit-shifting (PS) effect. Differentiating (12) with respect to the tariff imposed on imports from country $l$, $\tau_{il}$, yields $\forall\ l \neq i$ [1],

$$\frac{dW_i}{d\tau_{il}} = \underbrace{-\sum_{\substack{j=1\\ j\neq i}}^{n} q_{ij}\frac{dp^*_{ij}}{d\tau_{il}}}_{ToT\geq 0} + \underbrace{\sum_{\substack{j=1\\ j\neq i}}^{n} \tau_{ij}\frac{dq_{ij}}{d\tau_{il}}}_{VoT\leq 0} + \underbrace{(p_{ii}-w)\frac{dq_{ii}}{d\tau_{il}}}_{PS\geq 0}, \tag{13}$$

where $p^*_{ij}$ is the net-of-tariff price of country $j$'s good sold in country $i$, or $p^*_{ij} = p_{ij} - \tau_{ij}$. The ToT effect is the variation in the net-of-tariff price which country $j$'s firm receive for their exports to country $i$. In this model, the ToT effect is positive such that an increase in country $i$'s import tariff improves its terms of trade. The tariff reduces country $i$'s volume of trade ($VoT \leq 0$) due to a higher consumer price of imports, but it shifts profits from foreign exporters to domestic producers by reducing market access ($PS \geq 0$). This last effect is due to the oligopolistic distortion where the import tariff moves the domestic firm towards the Stackelberg leader output level. This effect would be absent under perfect competition where prices equal marginal costs. Moreover, if there were no strategic interaction between firms ($\gamma = 0$), then $\frac{dq_{ii}}{d\tau_{ij}} = 0$, such that there would be no profit-shifting incentive for imposing import tariffs. In this case, the only motive to unilaterally impose tariffs is to switch the terms of trade in its favour. However, when oligopoly matters there are two motives: to improve the terms of trade and to shift profits towards domestic firms. Substituting the Cournot quantities (5) and the inverse demand function (2) into (13), I can solve for the optimal non-cooperative tariff of country $i$ on imports from country $l$ [2]:

$$\tau^N_{il} = \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)A_i - D(1,n,\gamma)\alpha_{il}}{[\Gamma(0,\gamma)+1]D(1,n,\gamma)} > 0 \quad l \neq i, \tag{14}$$

where the expression $D(\cdot)$ is defined as:

$$D(k,n,\gamma) \equiv \Psi(k,\gamma)\Gamma(n,\gamma) + \Gamma(k,\gamma)\Gamma(2k,\gamma) \quad \text{and} \quad \Psi(k,\gamma) \equiv (\Gamma(0,\gamma)+1)\Gamma(k,\gamma) - \Gamma(2k,\gamma). \tag{15}$$

The superscript $N$ stands for Nash, and is there to illustrate the prisoner's dilemma nature of non-cooperative trade policy. Notice also from Appendix A that the governments do not set Nash tariffs strategically such that the Nash tariffs in country $i$ are independent of the Nash tariffs in other countries. This feature of the model is due to the assumption of market segmentation.

It is useful to see how changes in trade costs between a pair $i$ and $l$ affect country $i$'s incentive to impose a tariff on imports from country $l$. Taking the derivative with respect to $\alpha_{il}$ in (13) yields [3]:

$$\frac{d^2W_i}{d\tau_{il}d\alpha_{il}} = \frac{\gamma[\Gamma(2,\gamma)+\Gamma(n,\gamma)] - \Gamma(n,\gamma)^2}{[\Gamma(0,\gamma)\Gamma(n,\gamma)]^2} < 0. \tag{16}$$

Hence, when the cost of importing from country $l$, $\alpha_{il}$, falls, the gain from a tariff on imports from country $l$ increases. This is because when natural trade costs are lower, the natural distortion of profits and consumer prices is lower, making import tariffs more effective at switching the terms of trade in country $i$'s favour and shifting profits towards the domestic firm in the domestic market. Hence, the negative correlation between country $l$'s natural trade barrier and the optimal import tariff imposed on that country. Differentiating (14) yields:

$$\frac{d\tau^N_{il}}{d\alpha_{il}} = -\frac{1}{\Gamma(0,\gamma)+1} < 0.$$

[1] See Appendix A for the derivation.
[2] See Appendix A for the derivation.
[3] See Appendix A for the derivation.

The effect of natural trade costs of country $i$ and any other country than $l$ is either zero or positive. In fact, differentiating (14) with respect to trade cost from any other country, $\alpha_{ih}$, yields:

$$\frac{\tau_{il}^N}{d\alpha_{ih}} = \frac{\gamma\Gamma(0,\gamma)\Gamma(2,\gamma)}{[\Gamma(0,\gamma)+1]D(1,n,\gamma)} \geq 0, \qquad h \neq l.$$

The intuition is that the optimal import tariff on country $l$ compensates for trade diversion which results from increases in natural trade costs from other countries. However, when goods are independent of each other, $\gamma = 0$, there is no such diversion, and the relationship vanishes.

## 3.2 Cooperative trade policy

Unilateral trade policy is inefficient, however, as one country's welfare gain comes at the expense of the other. By taking the derivative of country $i$'s welfare with respect to country $l$'s import tariff on imports from country $i$, $\tau_{li}$, it is similarly possible to decompose the welfare effect into a terms-of-trade (ToT), a volume-of-trade (VoT) and a profit-shifting (PS) component[4]:

$$\frac{dW_i}{d\tau_{li}} = \underbrace{q_{li}\frac{dp_{li}^*}{d\tau_{li}}}_{ToT\leq 0} - \underbrace{\tau_{li}\frac{dq_{li}}{d\tau_{li}}}_{VoT\geq 0} + \underbrace{(p_{li} - w - \alpha_{li})\frac{dq_{li}}{d\tau_{li}}}_{PS\leq 0}, \tag{17}$$

where $p_{li}^*$ is the net-of-tariff price of firm $i$'s good sold in country $l$. Notice that country $l$'s import tariff only affects country $i$'s welfare through its export profits to that country. This is because of the assumption of market segmentation. As is clear from Appendix B, the derivative in (17) depends on country $l$'s Nash tariff on country $i$ as well as the sum of its tariffs imposed on all other countries, $T_l = \sum_{j=1}^n \tau_{lj}$. Evaluating (17) at country $l$'s Nash tariffs yields[5]:

$$\left.\frac{dW_i}{d\tau_{li}}\right|_{\tau_{li}^N, T_l} = \frac{\Gamma(1,\gamma)[\Gamma(n,\gamma)-\gamma]\Gamma(n,\gamma)\Gamma(0,\gamma)\gamma\Psi(1,\gamma)A_l + [\Gamma(0,\gamma)+1]\Psi(1,\gamma) - D(1,n,\gamma)\alpha_{li}]}{[\Gamma(0,\gamma)\Gamma(2,\gamma)]^2[\Gamma(0,\gamma)+1]D(1,n,\gamma)} < 0, \tag{18}$$

where $\Psi(\cdot)$ and $D(\cdot)$ are defined in (15). Hence, by acting non-cooperatively, international trade policy produces a terms-of-trade and a profit-shifting externality. It is also clear from (18) that this externality becomes more severe when trade costs are lower:

$$\left.\frac{d^2W_i}{d\tau_{li}d\alpha_{li}}\right|_{\tau_{li}^N, T_l} = \frac{\Gamma(1,\gamma)[\Gamma(n,\gamma)-\gamma][\Psi(1,\gamma)\Gamma(0,\gamma)(\Gamma(n,\gamma)-\gamma) - \Gamma(1,\gamma)\Gamma(0,\gamma)\Gamma(2,\gamma)]}{[\Gamma(0,\gamma)\Gamma(2,\gamma)]^2[\Gamma(0,\gamma)+1]D(1,n,\gamma)} > 0. \tag{19}$$

This feature of the model is not surprising. When trade costs fall, the unilateral gain from import tariffs is larger, and the Nash tariffs are therefore higher. This, in turn produces larger international externalities.

Under what circumstances would it improve global welfare to abandon import tariffs among all countries or a subset of countries? This is the question which will be addressed in the next two sections. For now, I will define what form cooperative trade policy takes. The countries in the world may decide to abandon import tariffs all together, and let trade flow freely amongst each other. Alternatively, a subset of the countries may form a trade agreement, and keep imposing positive Nash tariffs on all other countries not part of the agreement. Such a trade agreement may take the form of an FTA in which each country sets their own individually optimal external tariffs, or alternatively, they may form a CU in which they set common external tariffs. In the first case, assume that a subset of countries, $k$, sign a free trade agreement. Hence, if country $i$ is in $k$ it abolishes tariffs on all countries $1, ..., k$ but for any country $j = k+1, ..., n$ it sets the tariff which maximises its individual welfare. Each country $i$ in the FTA maximises the following:

$$\max_{\{\tau_{ij}\}_{j=k+1}^N} W_i = CS_i + TR_i + \Pi_i \tag{20}$$

Differentiating (20) with respect to the tariff imposed on imports from country $l$, $\tau_{il}$, yields $\forall\, l \neq i$ [6],

$$\frac{dW_i}{d\tau_{il}} = \underbrace{-\sum_{j=k+1}^{n} q_{ij}\frac{dp_{ij}^*}{d\tau_{il}}}_{ToT^{er}\geq 0} - \underbrace{\sum_{\substack{j=1\\j\neq i}}^{k} q_{ij}\frac{dp_{ij}}{d\tau_{il}}}_{ToT^{ir}\leq 0} + \underbrace{\sum_{j=k+1}^{n}\tau_{ij}\frac{dq_{ij}}{d\tau_{il}}}_{VoT^{er}\leq 0} + \underbrace{\sum_{\substack{j=1\\j\neq i}}^{k}\tau_{ij}\frac{dq_{ij}}{d\tau_{il}}}_{VoT^{ir}\geq 0} + \underbrace{(p_{ii}-w)\frac{dq_{ii}}{d\tau_{il}}}_{PS\geq 0}, \tag{21}$$

[4]See Appendix B for the derivation.
[5]See Appendix B for the derivation.
[6]See Appendix A for the derivation.

The optimal tariff balances country $i$'s *extra-regional* (er) terms of trade gain, $ToT^{er}$, against its *intra-regional* (ir) loss, $ToT^{ir}$. Substituting the Cournot quantities (5) and the inverse demand function (2) into (13), I can solve for the optimal extra-regional non-cooperative tariff of country $i$ on imports from country $l$ [7]:

$$\tau_{il}^{FTA} = \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)A_i - [D(k,n,\gamma) + \gamma^2\Gamma(1,\gamma)(k-1)(n-k)]\alpha_{il}}{D(k,n,\gamma) + \gamma^2\Gamma(1,\gamma)(k-1)(n-k)]\Gamma(0,\gamma)+1\}} \quad l \neq i. \tag{22}$$

The sign of the relationship between bilateral and total trade costs and country $i$'s import tariff, $\tau_{il}$ is the same as without the FTA, and the intuition is the same as above. It is clear from (22) that for $k = 1$ the optimal external tariff collapses into (14). In Appendix ?? I demonstrate that the optimal external tariff is larger than the Nash tariff. The intuition is that because of the FTA trade is diverted away from outside due to zero internal tariffs, and each country individually compensates by setting lower external tariffs. This is commonly referred to as the tariff complementarity effect in the trade policy literature. In the case of a CU, the countries in the union $1, ..., k$ find the external tariff which maximises their joint welfare:

$$\max_{\{\tau_{ij}\}_{i=1,j=k+1}^{k \quad N}} W_{CU} = \sum_{i=1}^{k} W_i = \sum_{i=1}^{k}\{CS_i + TR_i + \pi_{ii} + \sum_{\substack{j=1 \\ j\neq i}}^{n} \pi_{ji}\} \tag{23}$$

Differentiating (23) with respect to the import tariff of country $l$ yields:

$$\frac{dW_{CU}}{d\tau_l} = \underbrace{-\sum_{i=1}^{k}\sum_{j=k+1}^{n} q_{ij}\frac{dp_{ij}^*}{d\tau_l}}_{ToT\geq 0} + \underbrace{\sum_{i=1}^{k}\sum_{j=k+1}^{n} \tau_{ij}\frac{dq_{ij}}{d\tau_l}}_{VoT\leq 0} + \underbrace{\sum_{i=1}^{k}(p_{ii}-w)\frac{dq_{ii}}{d\tau_l}}_{PS\geq 0}. \tag{24}$$

Members of the customs union coordinate their internal terms of trade and profits, and effectively they act as if they were one country. Substituting the Cournot quantities (5) and the inverse demand function (2) into (13), I can solve for the optimal extra-regional tariff the CU on imports from country $l$ [8]: This gives the optimal external CU tariff as:

$$\tau_l^{CU} = \frac{\Gamma(0,\gamma)\Gamma(2k,\gamma)[\Gamma(0,\gamma)+1]k + \gamma^2(2k-n)[\Gamma(0,\gamma)+1]A_{CU}^{ir} + \gamma\Gamma(2k,\gamma)\Gamma(0,\gamma)A_{CU}^{er} - D(k,n,\gamma)A_l^{er}}{D(k,n,\gamma)[\Gamma(0,\gamma)+1]k}, \tag{25}$$

where $A_{CU}^{ir} \equiv \sum_{i=1}^{k} A_i$ is the aggregate *intra-regional* trade costs of the members of the customs union, $A_{CU}^{er} \equiv \sum_{i=k+1}^{n} A_i$ is their aggregate *extra-regional* trade costs, and $A_l^{er} \equiv \sum_{i=1}^{k} \alpha_{li}$ is the sum of country $l$'s bilateral trade costs with each member. Notice that the effect of $A_l^{er}$ on the common external tariff is unambiguously negative, $d\tau_l^{CU}/dA_l^{er} < 0$. The intuition for this is the same as for the one-country Nash tariff described above: when natural bilateral trade costs increase, an import tariff becomes less effective as a tool to manipulate the terms of trade and switch profits towards the CU market. The effect of the aggregate *extra-regional* trade costs on the import tariff depends on the degree of substitution which was also the case for the one-country Nash tariff. The optimal external tariff simply compensates for any trade diversion on other countries than $l$ if the substitutability between the goods is positive, $\gamma > 0$. The sign of the relationship between the aggregate intra-regional trade costs, $A_{CU}^{ir}$ depends on the size of the CU relative to the rest of the world. From (25) it is clear that $A_{CU}^{ir}$ affects the tariff on country $l$ positively whenever:

$$k > \frac{1}{2}n.$$

Hence, when the CU comprises more than half of the countries in the world, external tariffs will rise in response to increases in aggregate intra-regional trade costs. The intuition comes from the market size effect: when the CU is large, there is greater trade diversion within the CU when intra-regional trade costs rise. If the CU is small relative to the rest of the world this trade diversion is not very large. The tariffs in (14), (22) and (25) are ranked in the following lemma.

[7]See Appendix ??? for the derivation.
[8]See Appendix ??? for the derivation.

**Lemma 1** *The common external tariff in a customs union is strictly greater than the Nash tariff set by a country maximising its own welfare, and this Nash tariff is strictly greater than the common external tariff of an FTA. Proof. In Appendix ??* ∎

The common external tariff set by the CU is larger since its members are able to coordinate their internal profit-shifting and terms-of-trade manipulations as well as exploiting their increased monopoly power on the world market.

### 3.3 Self-enforcing trade policy

Following the convention in the trade policy literature, I model trade policy cooperation as a stationary repeated game, where cooperation is sustained only where it is incentive compatible for all countries. In this framework, each country weighs the benefit of defection against the future cost of such defection. I assume that countries can sustain cooperation through trigger strategies, where a defection by any country is followed by a permanent trade war.

In what follows it will be useful to define country $i$'s welfare expression in terms of two components, one which depends on surplus in the domestic market (i.e. consumer surplus, tariff revenue and local profits), and one which depends on producer surplus in all other countries. Denote the former by $S_i(T_i, A_i) = CS_i(T_i, A_i) + TR_i(T_i, A_i) + \pi_{ii}(T_i, A_i)$, which depends only on country $i$'s tariffs imposed on the rest of the world. The latter is given by $\sum_{\substack{j=1 \\ j\neq i}}^{n} \pi_{ji}(T_{-i}, A_{-i})$ which depends on the aggregate bilateral natural as well as political trade costs of the rest of the world less those of country $i$,

$$T_{-i} \equiv \sum_{\substack{j=1 \\ j\neq i}}^{n} T_j \quad \text{and} \quad A_{-i} \equiv \sum_{\substack{j=1 \\ j\neq i}}^{n} A_j.$$

Country $i$'s per-period welfare is thus:

$$W_i(T_i, T_{-i}, A_i, A_{-i}) = S_i(T_i, A_i) + \sum_{\substack{j=1 \\ j\neq i}}^{n} \pi_{ji}(T_{-i}, A_{-i}). \tag{26}$$

Assume that $k$ countries sign a trade agreement (TA) with one another, leaving aside the question of whether this is an FTA or a CU for the moment. Hence, $k$ countries abolish their tariffs on each other and if $k < n$ they will set external tariffs on any outsiders. Assume that country $i \in k$ and let $T_i^{TA} \equiv \sum_{j=1}^{n} \tau_{ij}^{TA}$ denote the sum of tariffs it imposes on other countries when it signs a TA, where $TA = FTA, CU$. This sum can be broken into its *intra-regional* and *extra-regional* components such that $T_i^{TA} = T_i^{ir} + T_i^{er} = \sum_{j=1}^{k} \tau_{ij} + \sum_{j=k+1}^{n} \tau_{ij}^{E} = \sum_{j=k+1}^{n} \tau_{ij}^{E}$, where $E = FTA, CU$. The per-period welfare from adhering to the trade agreement is:

$$\begin{aligned} W_i^{TA} = & \sum_{j=1}^{k} S_i(T_i^{TA}, A_i, \tau_{ij}^{TA}, \alpha_{ij}) + \sum_{j=k+1}^{n} S_i(T_i^{TA}, A_i, \tau_{ij}^{E}, \alpha_{ij}) + \sum_{\substack{j=1 \\ j\neq i}}^{k} \pi_{ji}(T_j^{TA}, A_j, \tau_{ji}^{TA}, \alpha_{ji}) \\ & + \sum_{j=k+1}^{n} \pi_{ji}(T_j^{N}, A_j, \tau_{ji}^{N}, \alpha_{ji}), \end{aligned} \tag{27}$$

where $T_j^N$ is the sum of Nash tariffs imposed by non-members. If country $i$ defects from the agreement, and reverts to its Nash tariff, it enjoys temporary benefits in the form of improved terms of trade and higher domestic profits at the expense of the members of the TA. In this case, its per-period welfare becomes:

$$W_i^{D} = \sum_{j=1}^{n} S_i(T_i^{N}, A_i, \tau_{ij}^{N}, \alpha_{ij}) + \sum_{\substack{j=1 \\ j\neq i}}^{k} \pi_{ji}(T_j^{TA}, A_j, \tau_{ji}^{TA}, \alpha_{ji}) + \sum_{j=k+1}^{n} \pi_{ji}(T_j^{N}, A_j, \tau_{ji}^{N}, \alpha_{ji}). \tag{28}$$

Notice that the one-shot defection welfare affects country $i$'s welfare only through domestic surplus, leaving its export profits unchanged. One period after defection, however, the other members of the agreement will punish the cheating nation by increasing their tariffs. It is not clear what this punishment might entail and in this paper I will consider two cases: one where deviation by one country implies contagion, such that every country in the agreement would revert to Nash against all members, and another where deviation does not imply contagion, such that only the cheating nation is punished and the rest of the countries would remain in a trade agreement. The per-period welfare of country $i$ following punishment by trading partners is:

$$W_i^P = \sum_{j=1}^{n} S_i(T_i^N, \tau_{ij}^N, \alpha_{ij}) + \sum_{\substack{j=1 \\ j\neq i}}^{k} \pi_{ji}(T_j^P, \tau_{ji}^P, \alpha_{ji}) + \sum_{j=k+1}^{n} \pi_{ji}(T_j^N, A_j, \tau_{ji}^N, \alpha_{ji}), \tag{29}$$

where $P = N, FTA, CU$. Country $i$ will stay in the agreement provided that the present discounted value of doing so is greater than or equal to the present discount value of deviation. It is then possible to define the following self-enforcement constraint:

$$\frac{1}{1+\delta} W_i^{TA} \geq W_i^D + \frac{\delta}{1+\delta} W_i^P \tag{30}$$

This constraint can be solved in terms of a critical $\delta_c$ above which the trade agreement can be sustained:

$$\delta_c \geq \frac{W_i^D - W_i^{TA}}{W_i^D - W_i^p} \tag{31}$$

The numerator of this expression is the one-shot gain in welfare from deviation whereas the denominator is the long-run cost. It will be convenient to define the one-shot benefit as:

$$\begin{aligned} B_i^D = W_i^D - W_i^{TA} &= \sum_{j=1}^{n} S_i(T_i^N, A_i, \tau_{ij}^N, \alpha_{ij}) - \sum_{j=1}^{k} S_i(T_i^{TA}, A_i, \tau_{ij}^{TA}, \alpha_{ij}) \\ &- \sum_{j=k+1}^{n} S_i(T_i^{TA}, A_i, \tau_{ij}^E, \alpha_{ij}). \end{aligned} \tag{32}$$

This benefit depends only on the change in domestic surplus. Similarly, I define the long-run cost of deviation as:

$$C_i^D = W_i^D - W_i^P = \sum_{\substack{j=1 \\ j\neq i}}^{k} \pi_{ji}(T_j^{TA}, A_j, \tau_{ji}^{TA}, \alpha_{ji}) - \sum_{\substack{j=1 \\ j\neq i}}^{k} \pi_{ji}(T_j^P, \tau_{ji}^P, \alpha_{ji}), \tag{33}$$

which depends on the change in surplus in other countries.

# 4 The two-country case

In this section I develop a simple model of two countries, call them country 1 and country 2. While this model does not address the issue of preferential trade agreements, it will develop the basic intuition which will serve as a starting point for the full triangular model presented in the next section. The purpose is to take an intermediate step towards establishing how reductions in natural trade costs may facilitate trade policy cooperation.

I choose country 1 as the home country, and due to symmetry of trade costs, $\alpha_{12} = \alpha_{21} = \alpha$ the analogous analysis for country 2 is a mirror image. Setting $n = 2$ in (14), it is easy to see that in a world with just two countries, Nash tariffs are given as:

$$\tau_{12}^N = \tau_{21}^N = \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)\alpha - D(1,2,\gamma)\alpha}{[\Gamma(0,\gamma)+1]D(1,2,\gamma)} \tag{34}$$

$$= \frac{1-\alpha}{3}. \tag{35}$$

Due to the assumed linear demand functions it is possible that the Nash tariffs become prohibitive. To rule this out I impose, for the two country case, the following condition on trade costs:

$$\alpha < \overline{\alpha} \equiv \frac{\Psi(1,\gamma)}{\Gamma(1,\gamma)^2} = 1 - \frac{3}{4}\gamma. \tag{36}$$

I establish this condition more formally in the following lemma:

**Lemma 2** *If and only if* $\alpha < \overline{\alpha} \equiv \frac{\Psi(\gamma)}{\Gamma(1,\gamma)^2}$*, there exists a unique non-prohibitive Nash tariff for both countries.*

Proof. See **??** ■

Notice that when the firms' monopoly power increases (a fall in $\gamma$), a larger trade cost can support a positive trade flow in equilibrium, that is, $\frac{d\overline{\alpha}}{d\gamma} = -\frac{3}{4} < 0$. This is because increased monopoly power allows the firms to produce at a higher mark-up, which in turn allows them to operate at larger marginal costs.

Acting unilaterally the two countries would enjoy benefits from deviation from an FTA in the form of improved terms of trade and higher profits in the domestic market. It was established in the previous section that this benefit is higher when natural trade costs are lower. Internationally, however, this implies that the externalities that tariffs impose are larger when trade costs are lower. Hence, when trade costs fall the conflict between unilateral and cooperative trade policy increases.

The next point which is necessary to examine is whether free trade improves joint welfare. Whenever goods are substitutable ($\gamma > 0$) this question is not certain. There are two effects that work in opposite directions: on the one hand, free trade brings joint benefits in the form of increased competition, on the other, trade implies a greater trade diversion since an exporting firm incurs wasteful trade costs[9]. This trade-off, however, is smaller when the substitutability between goods is lower (a lower $\gamma$). When $\gamma$ falls the gains from trade consist, to a larger extent, by the variety type, and thus the trade-off between pro-competition and trade diversion is smaller. When goods are independent of each other ($\gamma = 1$), there is no such trade off. In this case, the gain from trade arises exclusively from the increased variety, and therefore, free trade always improves joint welfare.

When $\gamma > 0$, it will be the case that for low trade costs the pro-competitive effect dominates the trade diverting effect of free trade, but as $\alpha$ exceeds a critical threshold, world welfare can be raised by increasing tariffs above zero to deal with the trade diversion. Let this critical threshold be denoted $\widehat{\alpha}$, such that from a bilateral perspective, tariffs should be raised above zero whenever:

$$\alpha > \widehat{\alpha} \equiv \frac{\Gamma(0,\gamma)^2}{\Gamma(2,\gamma)^2 - \gamma\Gamma(1,\gamma)^2} = \frac{(2-\gamma)^2}{4+\gamma^2}. \tag{37}$$

In the following lemma I state this result formally:

**Lemma 3** *If and only if* $\alpha > \widehat{\alpha} \equiv \frac{\Gamma(0,\gamma)^2}{\Gamma(2,\gamma)^2 - \gamma\Gamma(1,\gamma)}$*, bilateral welfare can be improved by raising tariffs above their free trade levels for the case* $\gamma > 0$*. Whenever* $\gamma = 0$ *free trade always improves welfare for any non-prohibitive trade cost.*
*Proof. See* **??** ■

Comparing (36) and (37) it can readily be established that $\overline{\alpha} > \widehat{\alpha}$ whenever $\gamma > 0$, whereas for $\gamma = 0$ they are equal, $\overline{\alpha} = \widehat{\alpha}$[10]. This result leads to the following lemma:

**Lemma 4** *For* $\gamma > 0$*, whenever trade costs are reduced by a negligible amount from their prohibitive levels, the trade-diverting effects of trade costs dominate the pro-competitive effect of free trade. In the case where* $\gamma = 0$*, such trade diversion does not occur and free trade always involves higher bilateral welfare than any positive tariff.*

In the simple two-country model it is easy to get an expression for the benefit from defection by evaluating (32) for $n = 2$. For country 1 this boils down to:

$$B_1^D = W_1^D - W_1^{TA} = S_1(\tau_{12}^N, \alpha) - S_1(\tau_{12}^{TA}, \alpha) = \frac{1-\alpha(2-\alpha)}{6\Gamma(0,\gamma)\Gamma(2,\gamma)}. \tag{38}$$

It is similarly possible to derive an expression for the long-run costs of defection by setting $n = 2$ in (33):

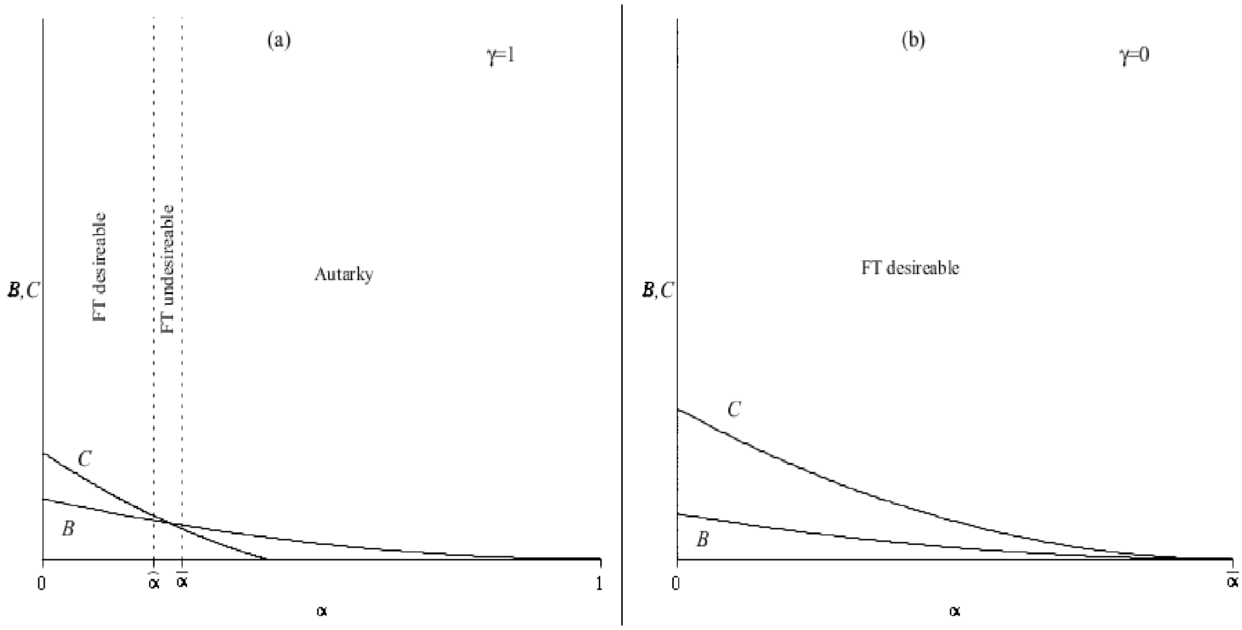
$$C_1^D = W_1^D - W_1^P = \pi_{21}(\tau_{21}^{TA}, \alpha) - \pi_{21}(\tau_{21}^N, \alpha) = \frac{5[1-\alpha(2-\alpha)] - 3\gamma(1-\alpha)}{[\Gamma(0,\gamma)\Gamma(2,\gamma)]^2}. \tag{39}$$

The relationship between the cost and benefits from defection and trade costs is established in the following proposition:

[9]This trade-off between pro-competition and trade diversion in the presence of trade costs was first established in Brander and Krugman (1983).

[10]Show this is this footnote

Figure 1: The cost and benefit of defection




**Proposition 5** *The benefit and cost of defection, respectively $B_1^D$ and $C_1^D$, are decreasing in $\alpha$ but positive for any non-prohibitive trade cost.*
*Proof. See Appendix ??* ■

The intuition is that when trade costs increase the natural distortion of consumer prices and profits is higher, making political trade costs less effective at doing the job. This implies that the unilateral incentive to deviate from an agreement decreases. The cost of deviation also decreases with $\alpha$. This is because when trade costs increase, the pro-competitive gains from free trade become lower relative to the losses in transit that results from higher trade costs. Hence, the Nash equilibrium becomes more attractive as trade costs increase. Hence, the fact that the benefits from deviation decreases in response to falling trade costs makes the TA more self-enforceable. However, since the costs of deviation also decrease, this makes the TA less enforceable. It is therefore necessary to analyse how the costs and benefits vary proportionally in response to changes in trade costs.

In Figure 1 I depict the costs and benefits of defection for the two extreme cases in which $\gamma = 1$ and $\gamma = 0$, respectively in panel (a) and (b). In panel (a) it is clear that the cost of deviation decreases proportionally faster than the benefits and at some critical value of the trade costs, call this $\tilde{\alpha}$, the benefits exceed the costs. In this case, a TA cannot be sustained. In panel (b), where I have set $\gamma = 0$ the costs always exceed the benefits for any non-prohibitive level of trade costs. The critical level of trade costs at which the benefits may exceed the costs of defection can be solved for from (38) and (39) as:

$$\tilde{\alpha} \equiv \frac{28 - 24\gamma + 3\gamma^2}{28 + 3\gamma^2}. \tag{40}$$

In the following proposition I establish the ranking of the critical levels of the $\alpha$s:

**Proposition 6** *For the case where $\gamma > 0$, the critical level at which a TA cannot be sustained due to the benefits exceeding the costs, $\tilde{\alpha}$, is smaller than the prohibitive level of the trade costs, $\overline{\alpha}$, but greater than the level at which free trade is no longer the desired bilateral trade policy, $\widehat{\alpha}$. Hence, the critical level of trade costs have the following ranking, $\overline{\alpha} > \tilde{\alpha} > \widehat{\alpha}$. For the case where $\gamma = 0$, they are equal such that $\overline{\alpha} = \tilde{\alpha} = \widehat{\alpha}$.*
*Proof. See Appendix ??* ■

The implication of this proposition is that when trade costs increase from zero, and when the degree of substitutability is greater than zero, the costs fall proportionally faster than the benefits from defection. On the other hand, when goods are independent of each other the costs and benefits decline in the same proportion in response to decreases in trade costs. The critical discount factor above which a TA is sustainable can be solved by diving the benefits of deviation in (38) by the costs in (39). This yields:

$$\delta_1^c = \frac{3(\Gamma(0,\gamma)\Gamma(2,\gamma)(1-\alpha)}{8(5(1-\alpha)-3\gamma)} \tag{41}$$

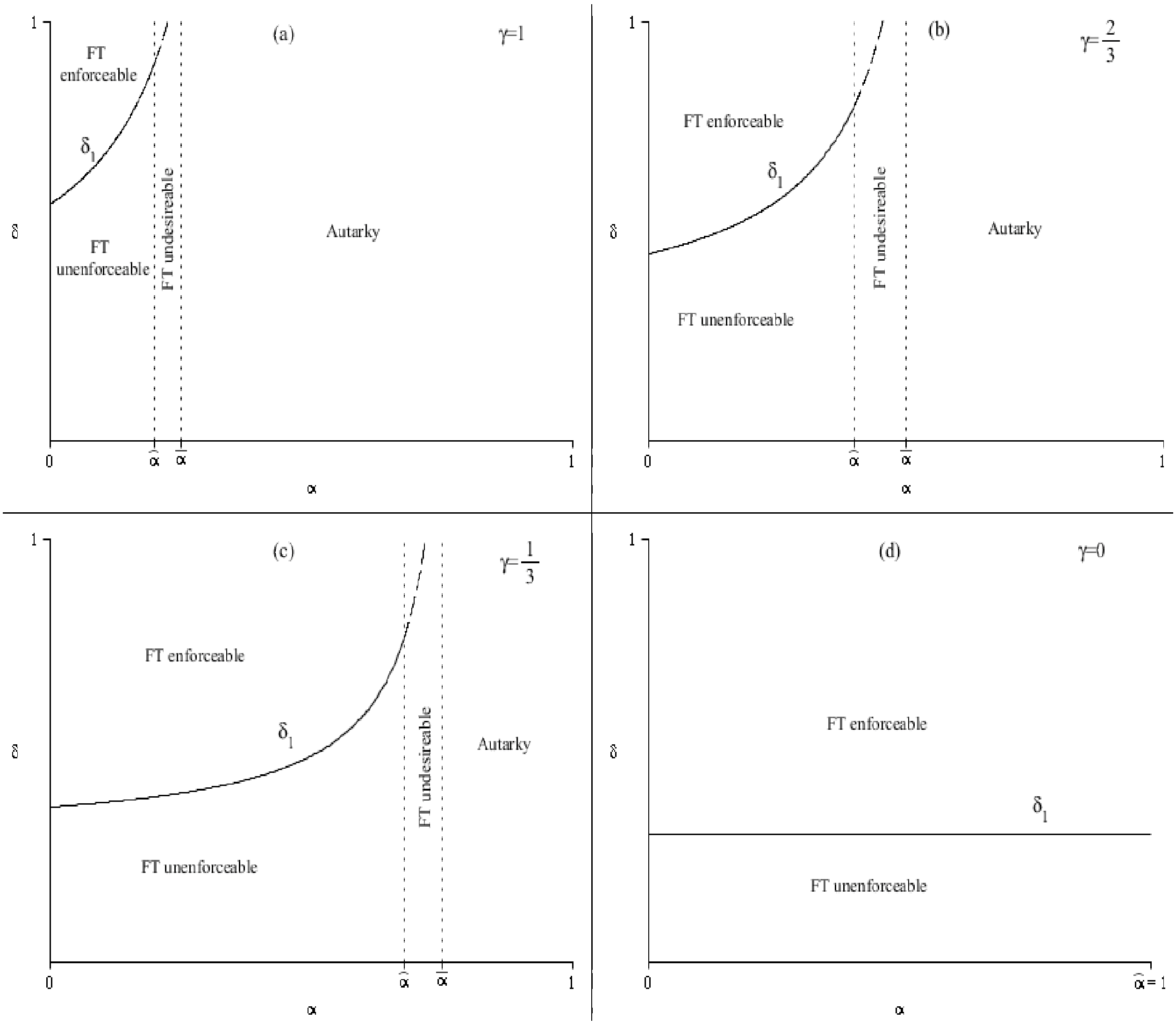
Differentiating $\delta_1^c$ with respect to $\alpha$ yields:

$$\frac{d\delta_1^c}{d\alpha} = \frac{9\gamma\Gamma(0,\gamma)\Gamma(2,\gamma)}{8(5(1-\alpha)-3\gamma)^2} \geq 0. \tag{42}$$

On the basis of (42) I can state the following proposition.

**Proposition 7** *For the case where $\gamma > 0$, decreases in trade costs make an FTA more self-enforceable by raising the critical discount factor, $\frac{d\delta_1^c}{d\alpha} > 0$. However, when $\gamma = 0$, the critical discount factor is unaffected by trade costs, $\frac{d\delta_1^c}{d\alpha} = 0$.*

Figure 2: The critical discount factors for different values of $\gamma$




In Figure 2 I depict the relationship between the critical discount factor and trade costs for four different values of $\gamma$. Starting in panel (a) it is clear that when goods are perfect substitutes, the range of non-prohibitive trade costs is smaller since greater competition reduces the markups of the two firms. Free trade is enforceable and desirable anywhere above the critical $\delta$ and in the range of trade costs $\alpha \in [0; \widehat{\alpha}]$. There is also a grey area in which free trade is enforceable but undesirable which is the range above the critical $\delta$ where it is depicted as a dashed line. This is the range where $\alpha \in [\widehat{\alpha}; \tilde{\alpha}]$. It can be seen from (42) that the relationship between $\delta_1^c$ and $\alpha$ becomes less steep when $\gamma$ is lower. Taking the derivative of (42) with respect to $\gamma$ yields,

$$\frac{d^2\delta_c(\alpha)}{d\alpha d\gamma} = \frac{3\left(20(1-\alpha) - 15\gamma^2(1-\alpha) + 12\gamma + 3\gamma^2\right)}{8\left(5(1-\alpha) - 3\gamma\right)^3} > 0.$$

This is also clear from panels (b) and (d) in Figure 2. When $\gamma$ falls, free trade can be supported for a larger range of discount factors, and the relationship between trade costs and the critical discount factor is less steep for every value of $\alpha$. The intuition is that when the competition between firms is less intense, the variety gain from trade is greater, implying that the pro-competition versus transit loss trade-off is less pronounced. This in turn means that free trade can be supported for a larger range of discount factors as well as a larger range of trade costs. The case of $\gamma = 0$ is depicted in panel (d), and it is clear that trade costs then become irrelevant for the sustainability of a TA in this two-country model. When the gains from trade arise exclusively from increased variety there is no trade-off between competition and losses in transit.

# 5 The full triangular model

In this section I will augment the simple two-country model with a third country, call this country 3. For convenience, I assume that the three countries are located on the corners of an isosceles triangle, with country 3 being farthest away in terms of trade costs but equidistant from its two trading partners. The purpose of the model is to examine how natural trade barriers affect the incentives of the three countries to engage in cooperative trade policy, whether this takes the form of global free trade or a preferential trade agreement. I assume that the trade costs between the two close partners is given as $\alpha_{12} = \alpha_{21} = \alpha_{ir}$ where the subscript *ir* stands for *intra-regional*. The trade cost between country 3 and 1, and that between country 3 and 2 are equal and given as $\alpha_{31} = \alpha_{13} = \alpha_{32} = \alpha_{23} = \alpha_{er}$ where the subscript *er* stands for *extra-regional*. For convenience I assume that intra-regional trade costs are zero, $\alpha_{ir} = 0$. With the additional structure of three countries, the Nash tariffs differ across countries. Due to symmetry of trade costs between country 1 and country 2, and the fact that they face the same extra-regional trade costs, the Nash tariffs which country 1 and 2 impose on each other and on the third country are equal. Setting $n = 3$ in (14), these tariffs can be written as:

$$\begin{aligned}\tau_{12}^N = \tau_{21}^N &= \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)(\alpha_{ir}+\alpha_{er}) - D(1,3,\gamma)\alpha_{ir}}{[\Gamma(0,\gamma)+1]D(1,3,\gamma)} \\ &= \frac{[\Gamma(0,\gamma)\Gamma(2,\gamma)][\Gamma(0,\gamma)+1] + \gamma[\Gamma(0,\gamma)\Gamma(2,\gamma)]\alpha_{er}}{[\Psi(1,\gamma)\Gamma(3,\gamma) + \Gamma(1,\gamma)\Gamma(2,\gamma)][\Gamma(0,\gamma)+1]}.\end{aligned} \tag{43}$$

Notice that this is increasing in $\alpha_{er}$. This is because greater extra-regional trade costs leads to increasing trade diversion. Each country inside the region compensate by setting higher intra-regional tariffs. The Nash tariff they impose on country 3 can also be found from (14):

$$\begin{aligned}\tau_{13}^N = \tau_{23}^N &= \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)(\alpha_{ir}+\alpha_{er}) - D(1,3,\gamma)\alpha_{er}}{[\Gamma(0,\gamma)+1]D(1,3,\gamma)} \\ &= \frac{[\Gamma(0,\gamma)\Gamma(2,\gamma)][\Gamma(0,\gamma)+1] + [\gamma[\Gamma(0,\gamma)\Gamma(2,\gamma)] - D(1,3,\gamma)]\alpha_{er}}{[\Psi(1,\gamma)\Gamma(3,\gamma) + \Gamma(1,\gamma)\Gamma(2,\gamma)][\Gamma(0,\gamma)+1]}.\end{aligned} \tag{44}$$

where the second lines in (43) and (44) use $\alpha_{ir} = 0$. It can readily be verified that (44) is decreasing in $\alpha_{er}$, since higher trade costs implies a smaller scope for terms of trade and profit-shifting manipulations. Country 3 faces the same trade costs regardless of its export destination so it will impose the same tariff on its two trading partner. These can similarly be found from (14):

$$\begin{aligned}\tau_{31}^N = \tau_{32}^N &= \frac{\Gamma(0,\gamma)\Gamma(2,\gamma)[\Gamma(0,\gamma)+1] + \gamma\Gamma(0,\gamma)\Gamma(2,\gamma)(\alpha_{er}+\alpha_{er}) - D(1,3,\gamma)\alpha_{er}}{[\Gamma(0,\gamma)+1]D(1,3,\gamma)} \\ &= \frac{\Gamma(0,\gamma)\Gamma(2,\gamma) - \Gamma(1,\gamma)[\Gamma(0,\gamma)\Gamma(2,\gamma) - \Gamma(1,\gamma)]\alpha_{er}}{2(3\Gamma(0,\gamma)[\Gamma(0,\gamma)-1] - \gamma)}\end{aligned} \tag{45}$$

This tariff rate is decreasing in $\alpha_{er}$. It is possible that the Nash tariffs prohibit trade between the two insiders and the outsider. However, since the two countries inside the region impose a different tariff on the outsider than the outsider impose on them it will be the case that the critical level of $\alpha_{er}$ which prohibits exports from the region to the outsider is different from the critical level of $\alpha_{er}$ which prohibits exports from the outsider to the region. Denote by $\overline{\alpha}_{r3}$, where $r = 1, 2$ is any of the two countries inside the region, the critical levels of trade costs below which country 3's export flows to the region, that is $q_{13}$ and $q_{23}$, are positive, and denote by $\overline{\alpha}_{3r}$ the critical levels of trade costs below which export flows to the outsider, that is $q_{31}$ and $q_{32}$, are positive. The critical levels can be solved as:

$$\alpha < \overline{\alpha}_{r3} \equiv \frac{\Psi(1,\gamma)[\Gamma(0,\gamma)+1]}{3\Gamma(0,\gamma)\Gamma(2,\gamma)} = \frac{(4-3\gamma)(3-\gamma)}{3(4-\gamma^2)} \quad \text{and} \tag{46}$$

$$\alpha < \overline{\alpha}_{3r} \equiv \frac{\Psi(1,\gamma)}{\Gamma(1,\gamma)^2} = 1 - \frac{3}{4}\gamma = 1 - \frac{3}{4}\gamma. \tag{47}$$

From inspection of (46) and (47), it is clear that $\overline{\alpha}_{r3} \leq \overline{\alpha}_{3r}$. Exports to the region are prohibitive at a lower level of trade costs for $\gamma > 0$, since consumers in country 1 and 2 are able to substitute for goods produced inside the region at a lower trade cost. When $\gamma = 0$ substitution cannot occur and the two critical thresholds are equal. In this paper I wish to focus on situations with trade in both directions such that the relevant condition to impose is that $\alpha < \overline{\alpha}_{r3}$.

Due to the trade-off between pro-competition and losses in transit, it will not necessarily be the case that it is in the three countries' long term interest to abolish political trade barriers between the region and the outsider. As in the two-country model, it is possible to derive critical levels of natural trade costs below which free trade raises welfare. However, those levels are different for a country inside the region relative to the outsider. Denote by $\widehat{\alpha}_3$ the critical level of trade costs below which free trade is raises welfare for the outsider. The critical threshold can be solved as:

$$\widehat{\alpha}_3 \equiv \frac{1}{2}\frac{[\gamma(0,\Gamma)+\gamma(0,1)][\gamma(0,\Gamma)+1]-\gamma}{\gamma\Gamma(2,\gamma)+\Gamma(1,\gamma)[\Gamma(2,\gamma)+1]} = \frac{1}{2}\frac{\gamma^2-8\gamma+12}{\gamma^2+4\gamma+6} \quad \text{and} \tag{48}$$

It is slightly more complicated to solve this for an insider, since such a country may derive a higher welfare from signing a preferential trade agreement with the other insider and leaving country 3 out. The question is: under what circumstances would it raise country 3's welfare to include the more distant country in a trade agreement? In Appendix ??, I prove that including country 3 in a trade agreement raises an insiders welfare provide the following condition is satisfied:

$$\alpha < \widehat{\alpha}_r \equiv \frac{3\Gamma(0,\gamma)\Gamma(2,\gamma) - 8\gamma\Gamma(0,\gamma)}{12 - \gamma\Psi(1,\gamma)}. \tag{49}$$

It is clear from (48) and (49) that global free trade raises welfare for an insider for a larger range of natural trade costs, $\widehat{\alpha}_3 \leq \widehat{\alpha}_r$ for $\gamma > 0$. This is not surprising, since free trade entails larger benefits for an insider since it is able to trade with its other inside partner at lower trade costs. When $\gamma = 0$ the two thresholds are equal to one, and it therefore always raises welfare to include country 3 in an agreement for every non-prohibitive trade cost.

The fact that intra-regional trade costs are assumed to be zero implies that it always raises welfare for one inside country to sign a preferential trade agreement with the other. In addition, if $\alpha < \widehat{\alpha}_r$ it would raise welfare for the insiders to include country 3 in an agreement. Each country has three options: (i) to remain in the Nash equilibrium, (ii) to sign a preferential trade agreement, or (iii) to sign a global free trade agreement. In this paper I allow two types of preferential trade agreements, an FTA and a CU. The next two subsections will carry out a welfare analysis of these two types.

## 5.1 The sustainability of an FTA versus global free trade

In order to determine the sustainability of PTAs it is necessary to examine whether a preferential FTA can occur in the presence of alternatives such as global free trade. The analysis of global free trade, however, poses further challenges since in this three country model, following deviation by one country, it is not clear what happens to the two remaining trading partners. As I mentioned above, in this paper I consider two scenarios. In the first, deviation by one country implies contagion such that in the following period all countries would return to the Nash equilibrium. In the second scenario the two countries that did not deviate would continue as a preferential trade agreement, leaving the defecting country out. I consider each scenario in turn.

#### 5.1.1 Contagion

To evaluate the sustainability of global free trade there are two constraints to consider: the critical discount factor required for the outside country to participate, and that which is required for any of the insiders. Setting $n = 3$ and $k = 3$ in (32) and (33) as well as evaluating them at their Nash tariff levels given in (44) and (43). These can thus be plugged into (31) and I can then write up the expression for the critical level of patience of an insider, say country 1 which by symmetry would come to the same value as that for country 2:

$$\delta_1^{GFT,C} = \frac{S_1(\tau_{12}^N, \tau_{13}^N, \alpha_{er}) - S_1(\tau_{12}^{FT}, \tau_{13}^{FT}, \alpha_{er})}{\pi_{21}(\tau_{21}^{FT}, \alpha_{er}) + \pi_{31}(\tau_{31}^{FT}, \alpha_{er}) - \pi_{21}(\tau_{21}^N, \alpha_{er}) - \pi_{31}(\tau_{31}^N, \alpha_{er})}, \tag{50}$$

where $GFT$ stands for global free trade, and $C$ stands for contagion. Similarly, the critical value for the outsider is:

$$\delta_3^{GFT,C} = \frac{S_3(\tau_{31}^N, \tau_{32}^N, \alpha_{er}) - S_3(\tau_{31}^{FT}, \tau_{32}^{FT}, \alpha_{er})}{\pi_{13}(\tau_{13}^{FT}, \alpha_{er}) + \pi_{23}(\tau_{23}^{FT}, \alpha_{er}) - \pi_{13}(\tau_{13}^N, \alpha_{er}) - \pi_{23}(\tau_{23}^N, \alpha_{er})}. \tag{51}$$

The relative magnitudes of these two critical values is established in the following lemma:

**Lemma 8** *For the case $\gamma > 0$ the critical discount factor required to sustain global free trade for the outsider is strictly greater than that for an insider, $\delta_3^{GFT,C} > \delta_1^{GFT,C}$. For the case $\gamma = 0$ they are equal, $\delta_3^{GFT,C} = \delta_1^{GFT,C}$.*
*Proof. See Appendix ??* ■

This result is not surprising since the inside countries are facing lower average trade costs with respect to their trading partners, and hence, their benefit of global free trade relative to their cost is lower permitting a larger range of discount factors to sustain free trade. The implication of the lemma is that the relevant constraint for global free trade is that facing the outside country, namely $\delta_3^{GFT,C}$.

I next examine the sustainability of a preferential trade agreement. There are two ways for such an agreements to occur: (i) the two insiders may sign a trade agreement, leaving country 3 out, or (ii) the distant country may sign a preferential trade agreement with one of the insiders. The external tariffs imposed by two countries signing an FTA can be found by setting $n = 3$ and $k = 2$ in (22). For country 1 or 2 this would be given as:

$$\tau_{13}^{FTA} = \tau_{23}^{FTA} = \frac{\Gamma(0,\gamma)\Gamma(2,\gamma) - \Gamma(2,\gamma)^2\alpha_{er}}{[12-\gamma^2][\Gamma(2,\gamma)-1] - \gamma^2\Gamma(2,\gamma)}, \tag{52}$$

whereas country 3 would impose the following tariff on either country 1 or 2:

$$\tau_{31}^{FTA} = \tau_{32}^{FTA} = \frac{\Gamma(0,\gamma)\Gamma(2,\gamma) - [\Gamma(0,\gamma)\Gamma(2,\gamma) - \gamma^2]\alpha_{er}}{[12-\gamma^2][\Gamma(2,\gamma)-1] - \gamma^2\Gamma(2,\gamma)}, \tag{53}$$

where the distant country's external tariff is larger since it faces a larger aggregate trade cost with respect to the rest of the world. The critical discount factor of the outsider which sustains a PTA with an insider, say country 1, can be found by setting $n = 3$ and $k = 2$ in (32) and (33) as well as using (53) and (52) and then plugging them into (31). This gives:

$$\delta_{31,-2}^{FTA} = \frac{S_3(\tau_{31}^N, \tau_{32}^N, \alpha_{er}) - S_3(\tau_{31}^{TA}, \tau_{32}^N, \alpha_{er})}{\pi_{13}(\tau_{13}^{TA}, \alpha_{er}) + \pi_{23}(\tau_{23}^N, \alpha_{er}) - \pi_{13}(\tau_{13}^N, \alpha_{er}) - \pi_{23}(\tau_{23}^N, \alpha_{er})}, \tag{54}$$

which by symmetry is equal to $\delta_{32,-1}^{FTA}$ The relative magnitudes of the outsiders discount factors of global free trade and for an FTA, respectively, is established in the following lemma:

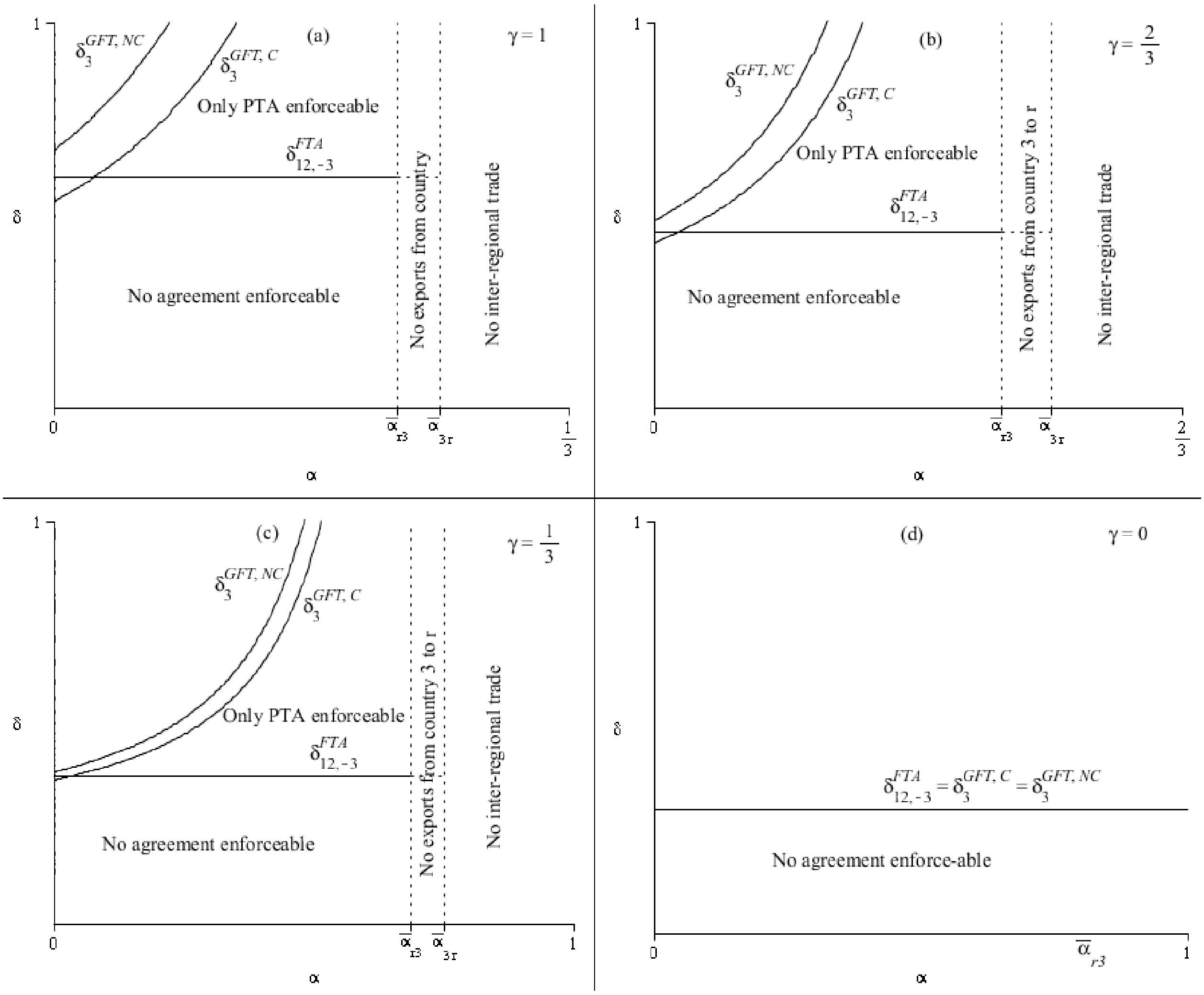
**Lemma 9** *For the case $\gamma > 0$ and with contagion, the critical discount factor required to sustain an FTA for the outsider is strictly greater than that required to sustain global free trade, $\delta_{31}^{FTA} > \delta_3^{GFT,C}$. For the case $\gamma = 0$ they are equal, $\delta_{31}^{FTA} = \delta_3^{GFT,C}$.*
*Proof. See Appendix ??* ■

The implication of this lemma is that if global free trade fails to be an equilibrium, then so does an FTA between distant countries. Hence, it is only necessary to consider the sustainability of an FTA between close partners and that of global free trade. The critical discount factor which is necessary to sustain an agreement between the close partners can be found by setting $n = 3$ and $k = 2$ in (32) and (33) as well as using (52) and (53) and then plugging them into (31). This gives:

$$\delta_{12,-3}^{FTA} = \frac{S_1(\tau_{12}^N, \tau_{13}^N, \alpha_{er}) - S_1(\tau_{12}^{TA}, \tau_{13}^N, \alpha_{er})}{\pi_{12}(\tau_{21}^{TA}, \alpha_{er}) + \pi_{31}(\tau_{23}^N, \alpha_{er}) - \pi_{21}(\tau_{21}^N, \alpha_{er}) - \pi_{31}(\tau_{31}^N, \alpha_{er})}, \tag{55}$$

and by symmetry this is equal to $\delta_{21,-3}^{FTA}$. The lower bound for the sustainability of global free trade is given in (51). It is now possible to plot these two critical values and this is done in Figure 3

Figure 3: The critical discount factors for different values of $\gamma$




For the case where defection implies contagion, global free trade is enforceable anywhere above the $\delta_3^{GFT,C}$-lines in panels (a) through (d). As in the two-country case,however, whenever $\gamma > 0$, global free trade is not enforceable for every non-prohibitive trade cost. Denote by $\tilde{\alpha}_3$ the level of trade costs, above which global trade is not enforceable for the pivotal country. In the following lemma I compare this level with the levels of trade costs that determine the desirability of free trade in (48) and (49).

**Lemma 10** *For $\gamma > 0$, $\tilde{\alpha}_3$ is strictly less than $\widehat{\alpha}_{r3}$ and $\widehat{\alpha}_{3r}$. For $\gamma = 0$ it is the case, $\tilde{\alpha}_3 = \widehat{\alpha}_{r3} = \widehat{\alpha}_{3r}$*
*Proof. In Appendix ??* ■

The implication of this lemma is that if free trade is enforceable, it is also desirable. This stands in contrast to the two-country case where there was a range of trade costs between $\widehat{\alpha}$ and $\tilde{\alpha}$ where free trade was enforceable but undesirable. The reason for the difference is as follows [state intuition for this finding]. It is clear from Figure 3 panel (a) that for low trade costs, global free trade can be supported for a larger range of discount factors, but when $\alpha$ exceeds a critical level, call this $\alpha_C^*$, there is scope for the forming a regional trade agreement. The same picture emerges in panels (b) and (d). In the following proposition I describe the relationship between the degree of product differentiation and the critical level of trade costs, $\alpha_C^*$.

**Proposition 11** *The critical level of trade costs, $\alpha_C^*$, above which there is scope for regionalism is increasing in $\gamma$, $\frac{\alpha_C^*}{d\gamma} > 0$.*
*Proof. In Appendix ??* ■

The implication of this proposition is that when products are more differentiated, and when the gains from trade to a larger extent are variety-driven, there is less scope for regionalism. In the special case where $\gamma = 0$ there is no scope for regionalism as any type of trade agreement, whether preferential or global are equally sustainable. In this case global free trade is desirable and if the world is patient enough to sustain global free trade this will happen. The next case to consider is the one where defection by one country does not imply contagion.

#### 5.1.2 No contagion

When there is no contagion, it will be the case that if country 3 deviates from global free trade the two close partners will remain in a preferential agreement. This does not alter the desirability or sustainability of a PTA in this three country framework but it will alter the incentive-compatibility of the pivotal country in a global FTA. The new constraint takes the following form:

$$\delta_3^{GFT,NC} = \frac{S_3(\tau_{31}^N, \tau_{32}^N, \alpha_{er}) - S_3(\tau_{31}^{FT}, \tau_{32}^{FT}, \alpha_{er})}{\pi_{13}(\tau_{13}^{FT}, \alpha_{er}) + \pi_{23}(\tau_{23}^{FT}, \alpha_{er}) - \pi_{13}(\tau_{13}^{FTA}, \alpha_{er}) - \pi_{23}(\tau_{23}^{FTA}, \alpha_{er})}. \tag{56}$$

There is only one subtle difference between (51) and (56) and that is that country 3 is punished by the optimal external tariff in an FTA $\tau^{FTA}$ rather than the Nash tariff. In lemma 1 I proved that the common external tariff chosen by member countries in an FTA is strictly lower than that which is chosen by countries individually due to the well-known tariff complementarity effect whereby each member compensates for the trade-diverting effects of the FTA by reducing tariffs on outsiders. Hence, for the case without contagion, country 3's punishment of defection is lower for every trade cost. In the next proposition I established the relationship between the sustainability of global free trade relative to a preferential FTA.

**Proposition 12** *For the case $\gamma > 0$, the critical discount factor of the pivotal country which sustains global free trade is strictly above that which sustains a preferential FTA for every non-prohibitive trade cost, $\delta_3^{GFT,NC} > \delta_{12,-3}^{FTA}$. For the case where $\gamma = 0$ they are equal, $\delta_3^{GFT,NC} = \delta_{12,-3}^{FTA}$.*
*Proof. In Appendix ??* ■

The implication of this proposition is that without contagion there is scope for regionalism for every level of trade costs. This finding is depicted in Figure 3. Whenever $\gamma > 0$, $\delta_3^{GFT,NC}$ is strictly above $\delta_{12,-3}^{FTA}$ in panels (a) through (c) and in the case $\gamma = 0$ all critical discount factors are the same.

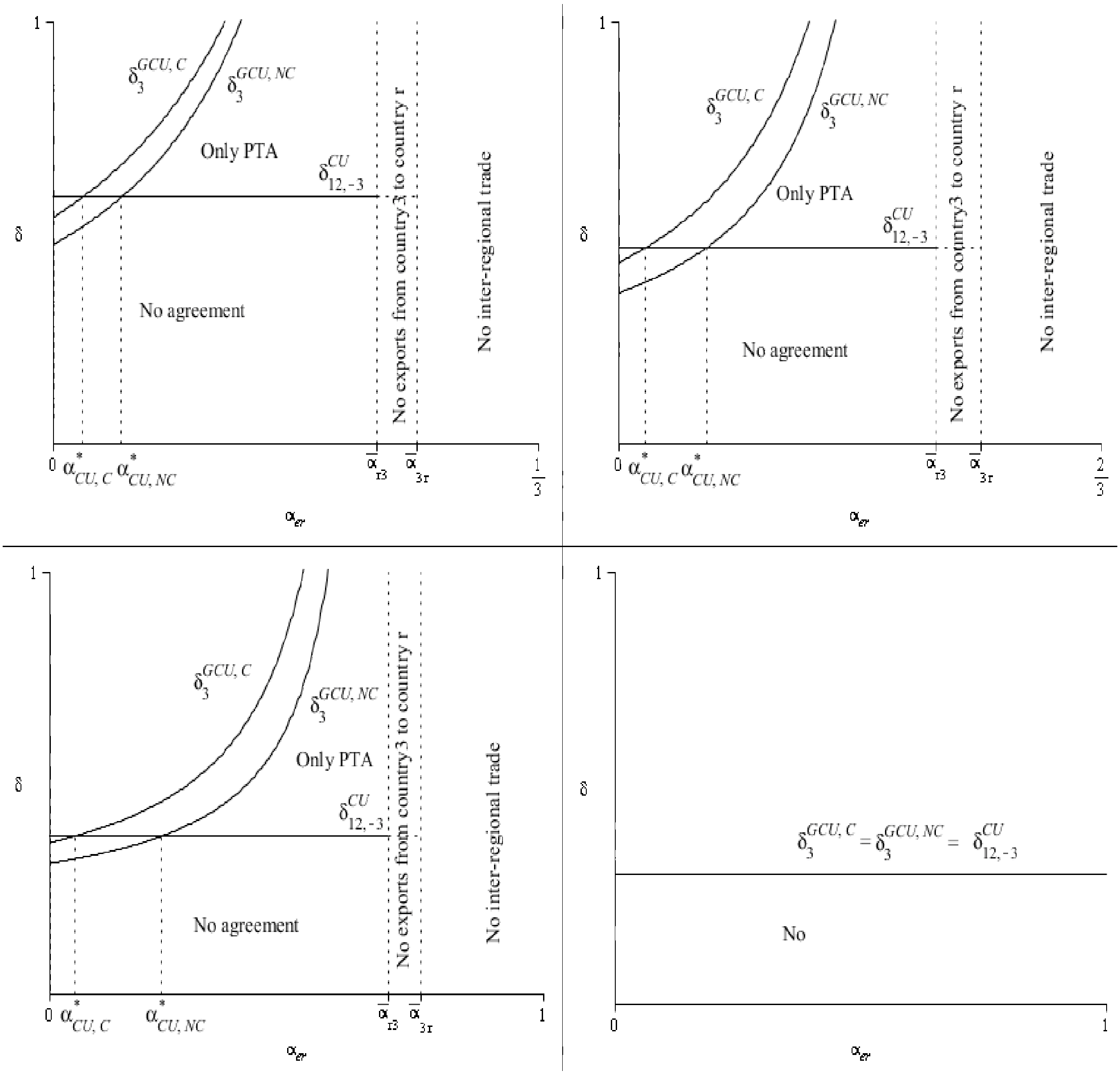
## 5.2 The sustainability of a Customs Union

In the case of a customs union, countries inside the union set a common external tariff on imports from outside. In doing so they coordinate their internal terms-of-trade and profit shifting effects, and together they will exert a larger market power on the outsider. Also, in this case it is necessary to assess whether a customs union can occur in the presence of an alternative such as global free trade. In the case of contagion, the critical discount factors for sustaining a global free trade agreement for any of the two close trading partners and for the distant country are the same as in (50) and (51). In the no contagion case, however, the fact that the non-deviating countries would remain as a customs union rather than a free trade agreement has important implications for the incentives to cooperate. Consider this from the point of view of the

$$\delta_3^{GCU,NC} = \frac{S_3(\tau_{31}^N, \tau_{32}^N, \alpha_{er}) - S_3(\tau_{31}^{FT}, \tau_{32}^{FT}, \alpha_{er})}{\pi_{13}(\tau_{13}^{FT}, \alpha_{er}) + \pi_{23}(\tau_{23}^{FT}, \alpha_{er}) - \pi_{13}(\tau_{13}^{CU}, \alpha_{er}) - \pi_{23}(\tau_{23}^{CU}, \alpha_{er})}. \tag{57}$$

Notice that the numerator of this expression is the same as in (56) but the denominator depends on the optimal external CU tariff rather than the optimal FTA tariff. This implies that the punishment of deviation is substantially larger than if the two non-defecting countries had punished the defecting nation by the FTA tariff. Hence, while the benefit of deviation remains unchanged the costs as larger which for every trade cost decreases the critical discount factor. In Figure 4 I depict the critical discount factors for the sustainability of global free trade and of a customs union.

Figure 4: The critical discount factors for different values of $\gamma$




State results for the customs union case as well. Use propositions.

# 6 Concluding remarks

Conclusion is being written.

APEPENDIX IS BEING COMPLETED

# Appendix A

## Derivation of Eq. (13)

Substituting (8), (9), (10) and (11) into (12) gives:

$$W_i = \sum_{j=1}^{n} \frac{1}{2}(a - p_{ij})q_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \tau_{ij} q_{ij} + (p_{ii} - w)q_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^{n} (p_{ji} - w - \tau_{ji} - \alpha_{ji})q_{ji} + wL_i. \tag{58}$$

Taking the derivative wrt. $\tau_{il}$, and noting that country $i$'s tariff does not affect firm $i$'s production decisions in other countries, the fixed wage rate, $w$, or the labour supply of country $i$, I obtain:

$$\frac{dW_i}{d\tau_{il}} = -\sum_{j=1}^{n} \frac{1}{2} \frac{dp_{ij}}{d\tau_{il}} q_{ij} + \sum_{j=1}^{n} \frac{1}{2}(a - p_{ij}) \frac{dq_{ij}}{d\tau_{il}} + \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \tau_{ij} \frac{dq_{ij}}{d\tau_{il}} + \frac{dp_{ii}}{d\tau_{il}} q_{ii} + (p_{ii} - w) \frac{dq_{ii}}{d\tau_{il}}. \tag{59}$$

Differentiating country $i$'s inverse demand function for firm $l$'s good (2) yields:

$$\frac{dp_{ij}}{d\tau_{il}} = -\frac{dq_{ij}}{d\tau_{il}} - \gamma \sum_{\substack{h=1 \\ h \neq j}}^{n} \frac{dq_{ih}}{d\tau_{il}}. \tag{60}$$

Substituting (2) and (60) into (59), noting that $\sum_{j=1}^{n} q_{ij} \sum_{\substack{h=1 \\ h \neq j}}^{n} \frac{dq_{ih}}{d\tau_{il}} = \sum_{j=1}^{n} \frac{dq_{ij}}{d\tau_{il}} \sum_{\substack{h=1 \\ h \neq j}}^{n} q_{ih}$, pure transfers cancel out:

$$\frac{dW_i}{d\tau_{il}} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} \frac{dp_{ij}}{d\tau_{il}} + q_{il} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \tau_{ij} \frac{dq_{ij}}{d\tau_{il}} + (p_{ii} - w) \frac{dq_{ii}}{d\tau_{il}}. \tag{61}$$

Defining the net-of-tariff price of firm $j$'s good sold in country $i$ as $p^*_{ij} = p_{ij} - \tau_{ij}$, I obtain the expression in (13). **q.e.d.**

## Derivation of Eq. (14) and (16)

Substituting the inverse demand function in (2), the Cournot quantites in (5) and their derivatives in (7) into (61) yields the following import-tariff first-order condition:

$$\begin{aligned}\frac{dW_i}{d\tau_{il}} = [\Gamma(0,\gamma)\Gamma(2,\gamma) + \gamma\{\Gamma(2,\gamma) + \Gamma(n,\gamma)[\Gamma(0,\gamma) + 1]\}T_i - \Gamma(n,\gamma)^2[\Gamma(0,\gamma) + 1]\tau_{il} \\ +\gamma\{\Gamma(2,\gamma) + \Gamma(n,\gamma)\}A_i - \Gamma(n,\gamma)^2\alpha_{il}]/[\Gamma(0,\gamma)\Gamma(n,\gamma)]^2 = 0,\end{aligned} \tag{62}$$

Summing this equation over all $n-1$ tariffs and multiplying out $(\Gamma(0,\gamma)\Gamma(n,\gamma))^2$ yield:

$$(n-1)\Gamma(0,\gamma)\Gamma(2,\gamma) + [\gamma\Gamma(n,\gamma) - \Gamma(1,\gamma)\Gamma(2,\gamma)]A_i + [\Psi(1,\gamma)\Gamma(n,\gamma) + \Gamma(1,\gamma)\Gamma(2,\gamma)]T_i = 0, \tag{63}$$

Solving (63) in terms of $T_i$ yields:

$$T_i = \frac{(n-1)\Gamma(0,\gamma)\Gamma(2,\gamma) + [\gamma\Gamma(n,\gamma) - \Gamma(1,\gamma)\Gamma(2,\gamma)]A_i}{D(1,\gamma)}, \tag{64}$$

where $D(\cdot)$ and $\Psi(\cdot)$ are defined in (15). Now substituting (64) into (62) yields the solution in (14). Finally, differentiating (62) wrt. $\alpha_{il}$ yields the expression in (16). **q.e.d.**

# Appendix B

## Derivation of Eq. (17)

Country $l$'s tariff on imports from country $i$ only affects production decisions in country $l$. Hence, differentiating (58) wrt. $\tau_{li}$ gives:

$$\frac{dW_i}{d\tau_{li}} = \left[\frac{dp_{li}}{d\tau_{li}} - 1\right] q_{li} + (p_{li} - w - \tau_{li} - \alpha_{li}) \frac{dq_{li}}{d\tau_{li}} \tag{65}$$

Defining the net-of-tariff price of firm $i$'s good sold in country $l$ as $p^*_{li} = p_{li} - \tau_{li}$, and rearranging yields the expression in (17). **q.e.d.**

## Derivation of Eq. (18)

Plugging the inverse demand function in (2), the Cournot quantites in (5) and their derivatives in (7) into (65) yields the following expression for the effect of country $l$'s import tariff on country $i$'s welfare:

$$\frac{dW_i}{d\tau_{li}} = \frac{\Gamma(1,\gamma)[\gamma - \Gamma(n,\gamma)][\Gamma(0,\gamma) + \gamma(T_l + A_l) - \Gamma(n,\gamma)(\tau_{li} + \alpha_{li})]}{[\Gamma(0,\gamma)\Gamma(n,\gamma)]^2}. \tag{66}$$

Evaluating this expression at country $l$'s specific Nash tariff on country $i$ in (14) and its aggregate Nash tariffs in (64) yield the expression in (18). **q.e.d.**

# Appendix C

To find the effect of the Foreign tariff on Home welfare I take the derivative of (**??**) wrt. $\tau_f$. This yields:

$$\frac{dW_h}{d\tau_f} = \frac{dp_{fh}}{d\tau_f} q_{fh} - q_{fh} + (p_{fh} - c - \tau_f - \alpha_f) \frac{dq_{fh}}{d\tau_f}. \tag{67}$$

Defining the net-of-tariff price of the Home good sold in the Foreign market as $p^*_{fh} = p_{fh} - \tau_f$ I obtain the expression in the (**??**). **q.e.d.**

**Derivation of Eq. (14), Eq. (??) and Eq. (18)**

Substituting the inverse demand functions in (2) and the Cournot quantities in (5) into (13), yields the following expression:

$$\begin{aligned} \frac{dW_h}{d\tau_h} = & \frac{(\Gamma(0,\gamma))^2 - \Gamma(0,\gamma)\Gamma(2,\gamma)(\tau_h + \alpha_h)}{(\Gamma(0,\gamma)\Gamma(2,\gamma))^2} \\ & -\tau_h \frac{2}{\Gamma(0,\gamma)\Gamma(2,\gamma)} + \frac{\gamma}{\Gamma(0,\gamma)\Gamma(2,\gamma)} \\ & - \left[\gamma \frac{\Gamma(0,\gamma) - 2(\tau_h + \alpha_h)}{\Gamma(0,\gamma)\Gamma(2,\gamma)}\right] \frac{\gamma}{\Gamma(0,\gamma)\Gamma(2,\gamma)}. \end{aligned} \tag{68}$$

Setting this expression equal to zero and rearranging I obtain:

$$-3\Gamma(0,\gamma)\Gamma(2,\gamma)\tau_h - \Gamma(0,\gamma)\Gamma(2,\gamma)\alpha_h + \Gamma(0,\gamma)\Gamma(2,\gamma) = 0.$$

Solving for $\tau_h$ yields the expression in (14). By writing up demand functions, Cournot quantities, and the expression for total welfare in the Foreign market, I could carry out the exact same steps for Foreign and obtain the optimal tariff for that country as well. Taking the derivative of (68) wrt. $\alpha_h$, I obtain the expression in (**??**). **q.e.d.**

By writing up expressions for the inverse demand functions and Cournot quantities in the Foreign market, an expression for (67) can be found as:

$$\frac{dW_h}{d\tau_f} = -\frac{\Gamma(0,\gamma) - 2(\tau_f + \alpha_f)}{\Gamma(0,\gamma)\Gamma(2,\gamma)} \frac{4}{\Gamma(0,\gamma)\Gamma(2,\gamma)}. \tag{69}$$

Substituting the Foreign Nash tariff from (14) into (69) yields the expression in (18). **q.e.d.**

**Proof of Proposition 1**

I need to show that the imported quantity in Home is positive when Home implements the Nash tariff in (14). In other words, it is sufficient to show that (from (5)):

$$q_{hf}\left(\tau_h^N\right) = \frac{\Gamma(0,\gamma) - 2\left(\tau_h^N + \alpha_h\right)}{\Gamma(0,\gamma)\Gamma(2,\gamma)} > 0.$$

Substituting the Nash tariff from (14) yields:

$$\frac{\Gamma(0,\gamma) - 2\left(\frac{1-\alpha_h}{3} + \alpha_h\right)}{\Gamma(0,\gamma)\Gamma(2,\gamma)} > 0.$$

Solving for $\alpha_h$ gives the expression in the proposition. By expressing the imported Cournot quantity in the Foreign market I could find the equivalent condition for the Foreign market. **q.e.d.**

**Derivation of Eq. (??), Eq. (??) and Eq. (??)**

The joint welfare of Home and Foreign is given as:

$$\frac{1}{1-\delta} J\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right) = \frac{1}{1-\delta}\left(W_h\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right) + W_f\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right)\right). \tag{70}$$

The expression for welfare in Home is obtained by substituting (**??**) and (**??**) into the expression for Home welfare (**??**). The equivalent expression for Foreign welfare is easily obtained by writing up the equivalent welfare expression by exchanging $h$ and $f$ in (**??**). Substituting the expressions for Home and Foreign welfare into (70) yields:

$$\begin{aligned} J\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right) &= CS_h\left(\tau_h, \alpha_h\right) + TR_h\left(\tau_h, \alpha_h\right) + \Pi_h\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right) \\ &\quad + CS_f\left(\tau_f, \alpha_f\right) + TR_h\left(\tau_f, \alpha_f\right) + \Pi_f\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right) \\ &= \frac{1}{2}\left(a - p_{hh}\right) q_{hh} + \frac{1}{2}\left(a - p_{hf}\right) q_{hf} + \tau_h q_{hf} \\ &\quad + \left(p_{hh} - c\right) q_{hh} + \left(p_{fh} - c - \tau_f - \alpha_f\right) q_{fh} \\ &\quad + \frac{1}{2}\left(a - p_{ff}\right) q_{ff} + \frac{1}{2}\left(a - p_{fh}\right) q_{fh} + \tau_f q_{fh} \\ &\quad + \left(p_{ff} - c\right) q_{ff} + \left(p_{hf} - c - \tau_h - \alpha_h\right) q_{hf}. \end{aligned} \tag{71}$$

Differentiating (71) with respect to $\tau_h$ yields

$$\begin{aligned} \frac{dJ}{d\tau_h} &= -\frac{1}{2}\frac{dp_{hh}}{d\tau_h} q_{hh} + \frac{1}{2}\left(a - p_{hh}\right)\frac{dq_{hh}}{d\tau_h} - \frac{1}{2}\frac{dp_{hf}}{d\tau_h} q_{hf} + \frac{1}{2}\left(a - p_{hf}\right)\frac{dq_{hf}}{d\tau_h} \\ &\quad + q_{hf} + \tau_h \frac{dq_{hf}}{d\tau_h} + \frac{dp_{hh}}{d\tau_h} q_{hh} + \left(p_{hh} - c\right)\frac{dq_{hh}}{d\tau_h} + q_{hf}\frac{dp_{hf}}{d\tau_h} \\ &\quad - \left(\tau_h + \alpha_h\right)\frac{dq_{hf}}{d\tau_h} + \left(p_{hf} - c\right)\frac{dq_{hf}}{d\tau_h}. \end{aligned} \tag{72}$$

Substituting the inverse demand functions (2) and the equivalent demand functions for Foreign into (72) and performing several algebraic steps reduces (72) to:

$$\frac{dJ}{d\tau_h} = \left(p_{hh} - c\right)\frac{dq_{hh}}{d\tau_h} + \left(p_{hf} - c\right)\frac{dq_{hf}}{d\tau_h} - \alpha_h \frac{dq_{hf}}{d\tau_h}, \tag{73}$$

which is the expression in (**??**). Substituting the inverse demand functions (2) and the Cournot quantities (5) and the equivalent functions for Foreign into (73) yields:

$$\begin{aligned} \frac{J\left(\tau_h, \tau_f, \alpha_h, \alpha_f\right)}{d\tau_h} &= \frac{-2\Gamma(0,\gamma)\Gamma(0,\gamma) + \alpha_h\left(4+\gamma^2\right) + \alpha_f\left(4+\gamma^2\right) - \tau_h\left(8 - 6\gamma^2\right)}{\left(\Gamma(0,\gamma)\Gamma(2,\gamma)\right)^2} \\ &= \frac{-2\Gamma(0,\gamma)\Gamma(0,\gamma) + \alpha_h\left(4+\gamma^2\right) + \alpha_f\left(4+\gamma^2\right)}{\left(\Gamma(0,\gamma)\Gamma(2,\gamma)\right)^2} - \frac{\tau_h\left(8 - 6\gamma^2\right)}{\left(\Gamma(0,\gamma)\Gamma(2,\gamma)\right)^2}. \end{aligned} \tag{74}$$

Solving for $\tau_h$ yields:

$$\tau_h^E = \frac{-\Gamma(0,\gamma)\Gamma(0,\gamma) + \frac{1}{2}\alpha_h\left(4+\gamma^2\right) + \frac{1}{2}\alpha_f\left(4+\gamma^2\right)}{\left(4 - 3\gamma^2\right)}. \tag{75}$$

I could carry out the same steps to find $\tau_f^E$, in which case I would find, by symmetry, that it is the same as $\tau_h$.

For symmetric trade costs $\alpha_h = \alpha_f = \alpha$ it is easy to show that, in the limit, the efficient tariffs and the Nash tariffs are equal. Setting (75) equal to (14) I obtain:

$$\frac{-\Gamma(0,\gamma)\Gamma(0,\gamma)+\alpha\left(4+\gamma^2\right)}{(4-3\gamma^2)}=\frac{1-\alpha}{3}.$$

Solving for $\alpha$ yields $\alpha = 1-\frac{3}{4}\gamma$, which is the upper bound proposed in (**??**). **q.e.d.**

**Derivation of Eq. (??), Eq. (??) and Eq. (??)**

In order to derive the level of patience required for an agreement based on the efficient tariffs in (**??**) to be sustained for Home (and symmetrically for Foreign), I need to evaluate Home's welfare in the three cases where: (i) both countries cooperate, (ii) both play Nash, and (iii) Home deviates by playing Nash while Foreign cooperates. I can find the expressions for Home welfare by substituting the inverse demand functions (2) and the Cournot quantities (5) into (**??**), and then evaluate them at the various tariff levels.Thus, after substantial algebraic manipulations, I obtain:

$$\begin{aligned}
W_h^C\left(\tau_h^E,\tau_f^E,\alpha\right) &= \frac{7+4\alpha^2-8\alpha-6\gamma(1-\alpha)}{2(4-3\gamma^2)};\\
W_h^P\left(\tau_h^N,\tau_f^N,\alpha\right) &= \frac{188+18\gamma^3(1-\alpha)-120\gamma(1-\alpha)+80\alpha^2+24\gamma^2\alpha-12\gamma^2\alpha^2-21\gamma^2-160\alpha}{18(2-\gamma)^2(2+\gamma)^2};\\
W_h^D\left(\tau_h^N,\tau_f^E,\alpha\right) &= \frac{-54\gamma^2(1-\alpha)-72\alpha\gamma^4+36\alpha^2\gamma^4+63\gamma^4+288\gamma^3(1-\alpha)-192\gamma^2\left(1+\alpha^2\right)}{6\left(4-3\gamma^2\right)^2(4-\gamma^2)}\\
&\quad+\frac{384\gamma^2\alpha-672\gamma(1-\alpha)+448\alpha^2+592-896\alpha}{6\left(4-3\gamma^2\right)^2(4-\gamma^2)}.
\end{aligned}$$

Substituting these expressions into (**??**) yields (**??**). **q.e.d.**

The minimum self-enforceable tariffs $\left(\tau_h^M,\tau_f^M\right)$ can be found by solving (**??**). Recall that the minimum self-enforceable tariffs will be chosen when the self-enforcement constraints bind. I first define the following expressions:

$$\begin{aligned}
\Phi_1\left(\tau_h,\tau_f,\tau_h^N,\tau_f^N\right) &= W_h^D\left(\tau_h^N,\tau_f,\alpha\right)-W_h^C\left(\tau_h,\tau_f,\alpha\right)\\
&\quad-\delta\left(W_h^D\left(\tau_h^N,\tau_f,\alpha\right)-W_h^P\left(\tau_h^N,\tau_f^N,\alpha\right)\right);\\
\Phi_2\left(\tau_h,\tau_f,\tau_h^N,\tau_f^N\right) &= W_f^D\left(\tau_f^N,\tau_h,\alpha\right)-W_f^C\left(\tau_f,\tau_h,\alpha\right)\\
&\quad-\delta\left(W_f^D\left(\tau_f^N,\tau_h,\alpha\right)-W_f^P\left(\tau_f^N,\tau_h^N,\alpha\right)\right).
\end{aligned}$$

Next I solve (**??**) using the lagrange method:

$$\begin{aligned}
\Psi(\tau_h,\tau_f,\lambda_1,\lambda_2) &= \max_{\tau_h,\tau_f}\frac{1}{1-\delta}\left(W_h(\tau_h,\tau_f,\alpha)+W_f(\tau_h,\tau_f,\alpha)\right)\\
&\quad+\lambda_1\left[\Phi_1\left(\tau_h,\tau_f,\tau_h^N,\tau_f^N\right)\right]\\
&\quad+\lambda_2\left[\Phi_2\left(\tau_h,\tau_f,\tau_h^N,\tau_f^N\right)\right],
\end{aligned}$$

where $\lambda_1$ and $\lambda_2$, respectively, are the lagrange multipliers of Home's and Foreign's self-enforcement constraints. Differentiating wrt. $\tau_h$, $\tau_f$, $\lambda_1$ and $\lambda_2$ yields:

$$\begin{aligned}
\frac{d\Psi}{d\tau_h} &= \frac{-4-\lambda_1\gamma^2\alpha\delta+8\lambda_2\delta-4\lambda_1+4\lambda_1\alpha\delta+\lambda_1\gamma^2\delta}{(1+\delta)(2-\gamma)^2(2+\gamma)^2}\\
&\quad-\frac{\lambda_1\gamma^2\alpha+4\lambda_2\delta\gamma+4\lambda_2\delta^2\gamma+8\lambda_2\delta\alpha+8\lambda_2\delta^2\alpha}{(1+\delta)(2-\gamma)^2(2+\gamma)^2}\\
&\quad+\frac{4\lambda_1\alpha-4\lambda_1\delta+\lambda_1\gamma^2+8\lambda_2\delta^2-4\tau_h+4\gamma+4\alpha}{(1+\delta)(2-\gamma)^2(2+\gamma)^2}\\
&\quad-\frac{\gamma^2-\gamma^2\alpha+3\lambda_1\gamma^2\tau_h\delta-12\lambda_1\tau_h\delta+3\lambda_1\gamma^2\tau_h}{(1+\delta)(2-\gamma)^2(2+\gamma)^2}\\
&\quad+\frac{3\gamma^2\tau_h+12\lambda_1\tau_h-8\lambda_2\delta^2\tau_h}{(1+\delta)(2-\gamma)^2(2+\gamma)^2};
\end{aligned}$$

$$\frac{d\Psi}{d\tau_f} = \frac{-4-4\lambda_2\delta-\lambda_2\gamma^2\alpha\delta-4\lambda_2-8\lambda_1\alpha\delta+4\lambda_2\delta\alpha}{(1+\delta)(2-\gamma)^2(2+\gamma)^2} - \frac{4\lambda_1\delta\gamma+4\lambda_1\delta^2\gamma+8\lambda_1\delta^2\alpha-\lambda_2\gamma^2\delta+\lambda_2\gamma^2\alpha}{(1+\delta)(2-\gamma)^2(2+\gamma)^2} + \frac{8\lambda_1\delta+8\lambda_1\delta^2+4\lambda_2\alpha+\lambda_2\gamma^2-4\tau_f+4\gamma}{(1+\delta)(2-\gamma)^2(2+\gamma)^2} + \frac{4\alpha-\delta^2+\delta^2\alpha+3\gamma^2\tau_f-8\lambda_1\delta\tau_f-8\lambda_1\delta^2\tau_f}{(1+\delta)(2-\gamma)^2(2+\gamma)^2} + \frac{12\lambda_2\delta\tau_f-3\lambda_2\gamma^2\tau_f+12\lambda_2\tau_f-3\lambda_2\gamma^2\tau_f\delta}{(1+\delta)(2-\gamma)^2(2+\gamma)^2};$$

$$\frac{d\Psi}{d\lambda_1} = \frac{-40\delta+12+80\alpha\delta+24\gamma\delta+12\alpha^2-3\gamma^2\alpha^2}{18(2-\gamma)^2(2+\gamma)^2} - \frac{40\delta\alpha^2+24\delta\gamma\alpha+72\tau_h+24\alpha+3\gamma^2-6\gamma^2\alpha}{18(2-\gamma)^2(2+\gamma)^2} - \frac{27\gamma^2\tau_h^2-72\tau_h\alpha-108\tau_h-144\delta\tau_f+72\delta\tau_f^2}{18(2-\gamma)^2(2+\gamma)^2} + \frac{18\gamma^2\tau_h-18\gamma^2\alpha\tau_h-72\delta\gamma\tau_f-144\delta\alpha\tau_f}{18(2-\gamma)^2(2+\gamma)^2};$$

$$\frac{d\Psi}{d\lambda_2} = \frac{-40\delta+12+80\alpha\delta+24\gamma\delta+12\alpha^2-3\gamma^2\alpha^2}{18(2-\gamma)^2(2+\gamma)^2} - \frac{40\delta\alpha^2+24\delta\gamma\alpha+72\tau_f+24\alpha+3\gamma^2-6\gamma^2\alpha}{18(2-\gamma)^2(2+\gamma)^2} - \frac{27\gamma^2\tau_f^2-72\tau_f\alpha-108\tau_f-144\delta\tau_f+72\delta\tau_h^2}{18(2-\gamma)^2(2+\gamma)^2} + \frac{18\gamma^2\tau_f-18\gamma^2\alpha\tau_f-72\delta\gamma\tau_h-144\delta\alpha\tau_h}{18(2-\gamma)^2(2+\gamma)^2};$$

Solving these four equations in the four unknowns $\tau_h$, $\tau_f$, $\lambda_1$ and $\lambda_2$ yields:

$$\lambda_1 = \lambda_2 = \frac{16\delta+3\gamma^2-6\gamma^2\delta-12}{\delta(1+\delta)(12-8\delta-3\gamma^2)};$$
$$\tau_h = \tau_f = \frac{\alpha\left(40\delta+3\gamma^2-12\right)-40\delta-3\gamma^2+12+24\delta\gamma}{3(12-3\gamma^2-8\delta)}. \quad (76)$$

It is easy to show that the minimum enforceable tariffs are equal to the Nash tariffs in the limit. Setting (76) equal to (14) yields:

$$\frac{\alpha\left(40\delta+3\gamma^2-12\right)-40\delta-3\gamma^2+12+24\delta\gamma}{3(12-3\gamma^2-8\delta)} = \frac{1-\alpha}{3}.$$

Solving for $\alpha$ yields $\alpha = 1-\frac{3}{4}\gamma$, which is the uppoer bound proposed in (**??**). **q.e.d.**

**Effects of trade costs on $\delta_c$**

I will begin by showing that the effect of $\alpha_h$ on the short-run gain of deviating from free trade, $W_h^D\left(\tau_h^N,\tau_f^C,\alpha_h,\alpha_f\right) - W_h^C\left(\tau_h^C,\tau_f^C,\alpha_h,\alpha_f\right)$, is negative. Substituting (**??**) and (**??**) into (**??**) yields an expression for Home welfare. Evaluating this when Home plays Nash by imposing the Nash tariff in (14), and when Foreign cooperates by choosing free trade, I have an expression for $W_h^D\left(\tau_h^N,\tau_f^C,\alpha_h,\alpha_f\right)$. Evaluating Home welfare when both set tariffs to zero gives an expression for $W_h^C\left(\tau_h^C,\tau_f^C,\alpha_h,\alpha_f\right)$. Using the inverse demand functions (2) and the Cournot quantites (5) it is possible to obtain, after substantial algebraic steps, the following:

$$W_h^D\left(\tau_h^N,\tau_f^C,\alpha_h,\alpha_f\right) - W_h^C\left(\tau_h^C,\tau_f^C,\alpha_h,\alpha_f\right) = \frac{(1-\alpha_h)^2}{6(2-\gamma)(2+\gamma)}. \quad (77)$$

Differentiating wrt. $\alpha_h$ yields:

$$\frac{d\left(W_h^D - W_h^C\right)}{d\alpha_h} = -\frac{2\left(1-\alpha_h\right)\alpha_h}{6\left(2-\gamma\right)\left(2+\gamma\right)} < 0.$$

Since this derivative is negative I can deduce that $\frac{d\delta_c}{d\alpha_h} < 0$, which was claimed in Proposition 2. Next I show that the effect of $\alpha_f$ on the long-run loss from not adhering to the FTA is also negative. Following a similar procedure I obtain:

$$W_h^D\left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right) - W_h^P\left(\tau_h^N, \tau_f^N, \alpha_h, \alpha_f\right) = \frac{4\left(1-\alpha_f\right)\left(5\left(1-\alpha_f\right)-3\gamma\right)}{9\left(2-\gamma\right)^2\left(2+\gamma\right)^2}. \tag{78}$$

Differentiating wrt. $\alpha_f$ yields:

$$\frac{d\left(W_h^D - W_h^P\right)}{d\alpha_f} = -\frac{4\left(10\left(1-\alpha_f\right)-3\gamma\right)}{9\left(2-\gamma\right)^2\left(2+\gamma\right)^2} < 0.$$

Similarly, since this derivative is negative I can deduce that $\frac{d\delta_c}{d\alpha_f} > 0$, which was also claimed in Proposition 2.

Setting $\alpha_h = \alpha_f = \alpha$ and dividing (77) by (78) yields an expression for the critical discount factor for symmetric trade costs:

$$\delta_c\left(\alpha\right) = \frac{3\left(1-\alpha\right)\left(4-\gamma^2\right)}{8\left(5\left(1-\alpha\right)-3\gamma\right)}.$$

which is the expression in (**??**). Taking the derivative wrt. $\alpha$ yields:

$$\frac{d\delta_c}{d\alpha} = \frac{30\gamma\left(4-\gamma^2\right)}{8\left(5\left(1-\alpha\right)-3\gamma\right)^2} \geq 0.$$

It is clear that when $\gamma = 0$ this derivative is negative, and when $\gamma > 0$ it is strictly positive. This is what Proposition 3 claims. **q.e.d.**

# References


Bagwell, K. and Staiger, R. W. (2010). The world trade organization: Theory and practice. *Annual Review of Economics*, 2(1):223–256.

Baldwin, R. E. and Venables, A. J. (1995). Regional economic integration. In Grossman, G. M. and Rogoff, K., editors, *Handbook of International Economics*, volume 3 of *Handbook of International Economics*, chapter 31, pages 1597–1644. Elsevier.

Brander, J. and Krugman, P. (1983). A 'reciprocal dumping' model of international trade. *Journal of International Economics*, 15(3-4):313–321.

Brander, J. A. (1981). Intra-industry trade in identical commodities. *Journal of International Economics*, 11(1):1–14.

Freund, C. and Ornelas, E. (2010). Regional trade agreements. *Annual Review of Economics*, 2(1):139–166.

Ludema, R. D. (2002). Increasing returns, multinationals and geography of preferential trade agreements. *Journal of International Economics*, 56(2):329–358.

Mrazova, M. (2009). Trade negotiations when market access matters. Technical report.

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.