The University of Nottingham

# research paper series

**Theory and Methods**

Research Paper 2011/26

## *The Self-enforceability of Trade Agreements in the Presence of Trade Costs*

by

**Christian Soegaard**

The Centre acknowledges financial support from The Leverhulme Trust under Programme Grant F/00 114/AM

GEP Leverhulme Centre for Research on Globalisation and Economic Policy

**The Authors**

Christian Soegaard is a PhD student at the Leverhulme Centre for Research on Globalisation and Economic Policy at the University of Nottingham.

# The Self-enforceability of Trade Agreements in the Presence of Trade Costs

**by**

**Christian Soegaard**

**Abstract**

This paper sets up a two-country model of oligopoly to analyse the relationship between trade costs and trade policy cooperation. Acting non-cooperatively, the two 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 the other. The incentive to do this is higher when trade costs are lower. Cooperative trade policy, on the other hand, is concerned with minimising losses in transit, such that internationally efficient tariffs are lower when trade costs fall. Hence, there is a conflict of interest between unilateral and cooperative trade policy in response to reductions in trade costs. I then analyse trade policy cooperation which must be sustained by a reputational mechanism. I first demonstrate that, provided the two countries care sufficiently about the future, lower import tariffs are more self-enforceable when trade costs are lower. I also find that global free trade can be supported for a larger range of discount factors in response to falling trade costs, provided firms interact strategically.

**JEL classification:** F13, F15

**Keywords:** trade policy, self-enforceability, trade costs, free trade agreement.

**Outline**

**Non-Technical Summary**

In this paper, I construct a model of international oligopoly where two countries are unilaterally driven into a terms-of-trade and a profit-shifting prisoner's dilemma in the setting of trade policy. In addition to political import tariffs, world trade is subject to natural trade costs (or transport costs) which are outside the government's control.

I first identify key relationships between natural trade costs and trade policy. Unilaterally, the incentive to distort a country's terms of trade in its favour, and the incentive to shift profits towards its domestic market increases when trade costs decline. This is because when natural trade costs are lower, the degree of natural distortion of consumer prices, and the degree of natural profit-shifting towards the domestic market are lower, making import tariffs more effective at doing the job. Cooperative trade policy, however, has a different objective. Acting cooperatively the two countries set import tariffs to minimise global losses in transit. Hence, when trade costs are lower, a smaller amount of the traded quantities are lost in transit, facilitating lower cooperative tariffs. This implies that when trade costs decline, a conflict of interest between unilaterally optimal trade policy and cooperative trade policy intensifies.

I go on to examine the sustainability of cooperative trade policy when any trade policy must be sustained through a reputational mechanism. In an infinitely repeated game, I find that lower import tariffs can be lowered in response to falling trade costs provided the governments' long-run objectives of minimising losses in transit are sufficiently important relative to each government's short-run temptation to distort the terms of trade in its favour, and shifting profits towards the domestic market. In an impatient world, on the other hand, cooperative tariffs must be raised in response to lower trade costs, in order to keep each country's incentive to stay in the agreement.

I also show that the two countries can support lower import tariffs in response to falling trade costs for the largest range of discount factors. Moreover, I show that a free trade agreement can be supported for a larger range of discount factors when trade costs fall, provided the firms interact strategically. The intuition behind this is that although the temptation to deviate from a free trade agreement increases when trade costs decline, the international externalities that are felt by both countries when punishment occurs outweigh the short run gains.

# 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. 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? Did it occur because governments became aware of the harmfulness of non-cooperative trade policies, or was it caused by external events which made cooperation more favourable?

Along side the substantial reduction in politically-induced tariff protection, the post war era has witnessed a gradual decline in the overall level of trade costs. Figure 1 is taken from McGowan and Milner (2011) and it shows the trend in trade costs over the past three decades. Trade costs are measured using Novy's (2010) gravity approach, and it effectively gives a measure of the ratio of external trade barriers to internal barriers. In the figure trade costs, for a sample of developed countries, have been averaged to give an idea of the overall trade costs in the developed world. The data begins in 1980 and records average trade costs of a little under 450 percent of internal trade costs. In 2006, average trade costs have dropped to around 135 percent. In this paper, I study the nature of the relationship between trade costs and the apparent extensive trade liberalisation which has occurred up until now.

I construct a modified version of Yi's (1996) extension of the Brander (1981) model of oligopoly with two countries called Home and Foreign, and with one firm in each. Each firm produces one good, and these goods are substitutable according to a substitution index which varies between zero and one. If the index is zero each firm is a monopolist and there is no strategic interaction amongst them. If the index is one, the two firms produce homogeneous goods. Unilaterally, each country maximises individual welfare with respect to an import tariff, and the model offers two motives to grant such import protection: (i) a terms-of-trade motive, and (ii) a profit-shifting motive. The first motive arises due to the influence of import tariffs on the net-of-tariff price of an imported good. This motive is standard in the trade-policy literature, and it was first identified by Johnson (1954) for perfectly competitive

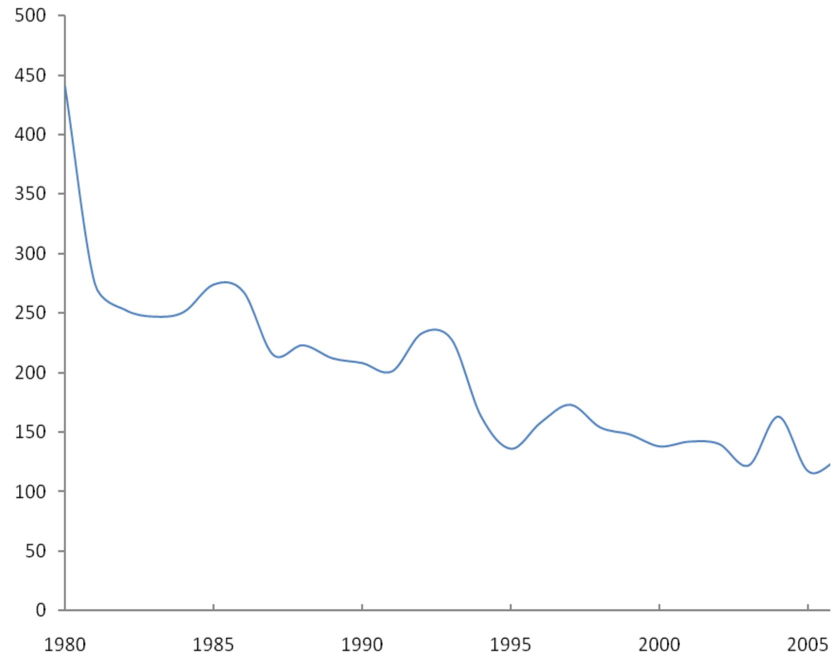


Figure 1: Evolution of trade costs over time.

economies[1]. The key is that a country as a whole can use its monopoly power over its terms of trade to improve welfare. The second motive arises due to rent-shifting made possible by the oligopolistic distortion. By reducing market access for the foreign exporter, import tariffs can shift profits towards the domestic producer. This motive is absent under perfect competition where prices are equal to marginal costs.

I make a clear distinction between trade costs that are politically induced which in the present model consist of import tariffs, and those trade costs which are natural in the sense that governments are unable to influence them through policy. In the trade literature, it is most common to refer to such trade barriers as transport costs, but I stress *natural* trade costs because I want these to comprise all kinds of barriers to trade which are outside the influence of government. These could include: transport technology, storage, inventory and preparation technology, communications networks, language barriers and so on.

The incentives to impose import tariffs through improved terms of trade and higher domestic profits are not invariant to changes in natural trade costs. In fact, when natural trade costs are lower the unilateral gain from imposing import tariffs is higher. This is because when Home's consumer prices and domestic profits are distorted to a lesser extent by natural trade costs, a politically induced import tariff is more effective at distorting terms

[1] Microfoundations for such a terms-of-trade distortion is modelled in Grossman and Helpman (1995).

of trade in Home's favour, and shifting profits towards the Home producer in the domestic market.

Unilateral trade policy produces international externalities, however, since a country's trading partner will see a worsening of its terms of trade and profits shifted away from their firm. These international distortions are greater the lower the level of natural distortions through trade costs. The countries engaged in a tariff war will find themselves caught in a terms-of-trade and a profit-shifting prisoner's dilemma and they may be willing to cooperate in order to reach a more efficient equilibrium.

I first consider an equilibrium in which Home and Foreign are able to sign a binding trade agreement which is perfectly enforceable. In this situation, Home and Foreign simply maximise joint welfare with respect to import tariffs to reach the optimal level of joint welfare, thus neutralising the externalities of unilateral trade policy. However, since the presence of natural trade costs represents losses in transit, globally efficient trade policy involves setting lower tariffs when trade costs fall. More specifically, due to the oligopolistic distortion firms produce suboptimal quantities and in the absence of trade costs, it is optimal to subsidise them. When trade costs are positive, however, the governments balance subsidising the oligopolistic firms against minimising losses in transit. Hence, when acting to maximise global welfare, the two countries will set lower tariffs in response to exogeneous decreases in natural trade costs. There is, therefore, a conflict of interest between unilateral and global trade policy: unilaterally, it is optimal to raise tariffs when trade costs fall, whereas bilaterally it is optimal to lower the tariffs. The joint welfare gain from cooperation is thus larger when trade costs are lower.

I then move on to discuss trade agreements in a framework where the two countries are unable to write binding contracts. While the literature on trade agreements offers several approaches to modelling imperfect contract enforcement, one particular approach which is based on repeated games stands out. According to these models, trade policy cooperation is limited by countries' weighting the one-shot incentives to deviate from an agreed-upon tariff against the discounted benefits from future cooperation. This view of trade policy cooperation is a fair reflection of reality since the world is currently not equipped with an international law enforcement agency capable of sanctioning nations that do not honour international agreements. The GATT/WTO has served as an international institution offering a negotiation forum, which allows countries to reach a higher welfare through mutual tariff concessions. Outside enforcement is ensured through a number of rules, permitting countries to punish cheating nations[2].

In this paper, I model trade policy cooperation as a repeated prisoner's dilemma. The two countries can choose to sign any agreement specifying that tariffs be lowered from their non-cooperative levels. Once such an agreement is signed, however, any of the two countries has a one-shot incentive to deviate from an agreed-upon tariff by distorting the terms of trade in their favour, and shifting profits towards their respective domestic firms when the other country cooperates. It is assumed that if a country does not honour an agreed-upon tariff, the other country will punish it by reverting to the non-cooperative Nash tariffs forever as of the following period. Self-enforceability thus implies that the present discounted value of honouring a trade agreement must be greater than or equal to the one-shot payoff from deviating and the present discounted value of infinite Nash reversion.The requirement of self-enforceability may, therefore, constrain the trade agreement to a second-best one from a joint welfare-maximising perspective in order to keep each country's incentive to stay in the agreement.

I consider two types of agreements which are subject to a self-enforcement constraint. The first type is based on what I define as the optimal self-enforceable tariff, and the second is a free trade agreement (FTA).

The first type is the tariff which obtains when the two countries maximise their joint welfare with respect to import tariffs subject to the self-enforcement constraints. If the world is very patient the self-enforcement constraint is not binding, and the two countries are simply able to implement the internationally efficient tariffs. On the other hand, if short-run considerations are so important that the the self-enforcement constraint binds, the agreement must involve tariffs that lie somewhere in between the efficient tariffs and the politically optimal Nash tariffs in order to persuade each country to adhere to the agreement. I then go on to examine how exogeneous changes in natural trade costs affect the optimal self-enforceable tariffs. It turns out that if countries are very impatient such that short-run considerations are important, tariffs need to be raised in response to falling trade costs. This is because when trade costs fall the one-shot payoff from deviating from an agreed-upon tariff increases, since there is a greater benefit from distorting terms of trade in a country's favour, and shifting profits towards the domestic firm . On other words, in order to keep an

[2] For example, Article 22.3 of the Understanding on Rules and Procedures Governing the Settlement of Disputes, and a limited punishment rule by the GATT Article XXVIII.

impatient country's incentive to stay in the agreement, tariffs need to be raised in response to falling trade costs in order to reduce the short-run payoff from unilateral deviation. If the world is sufficiently patient, however, the long run payoff from cooperation is relatively more important, and for that reason the joint objective of minimising losses in transit becomes more important. While this may not imply that the internationally efficient tariffs are self-enforceable, countries care sufficiently about the future to sustain lower tariffs in response to falling trade costs. Hence, the relationship between tariffs and trade costs depends crucially on how patient the world is. I demonstrate that tariffs fall in response to falling trade costs for the largest range of discount factors, allowing the two countries to increase their welfare.

The second type of agreement, a FTA, is useful for several reason. First, the post-war era has witnessed the emergence of a large number of FTAs, and it is therefore interesting to see how trade costs affect the sustainability of such an agreement. Second, because a FTA involves removing import tariffs as a trade policy instrument it may be that such an agreement is the most self-enforcing if there are fixed costs associated with reinstating customs and border controls. Although I do not model such fixed costs explicitly, they could easily be added to the model. I define a critical discount factor above which a FTA is self-enforceable. This discount factor is given as the ratio of the one-shot gain from deviating from a FTA to the long-run loss from infinite Nash reversion. I demonstrate that the one-shot gain from deviating is greater when trade costs fall. This is because, relative to free trade, the one-shot benefit from distorting the terms of trade and shifting profits towards the domestic firm is higher when the natural distortion due to trade costs is lower. The long run loss from infinite Nash reversion is also higher when trade costs fall, since the punishment in terms of lost profit in the export market and worsened terms of trade is higher when trade costs are lower. Provided that firms interact strategically, I demonstrate that although the one-shot benefit from deviating from a FTA is higher when trade costs are lower, the greater international externalities that would result from deviation ensures that a FTA is more sustainable when trade costs fall. Put differently, a FTA can be supported for a larger range of discount factors when trade costs are lower. This relationship is steeper when the degree of strategic interaction is higher. I also show that there is a discontinuity in the relationship between trade costs and trade policy cooperation. In fact, there is a trade cost threshold that must be crossed before a FTA can be supported at all. This is because, when trade costs are above this threshold there are too many losses in transit such that free trade is far too deep a trade liberalisation from a joint welfare-maximising perspective. If firms do not interact strategically, there is no relationship between the critical discount factor and trade costs. This is because, when each firm is a monopolist in its own market, the change (with respect to trade costs) in the unilateral gain from deviation is exactly proportional to change in the long run loss. Hence, there is no such relationship.

I would argue that this model is able to account for several aspects of the post-war era trade liberalisation. First, provided the world has been sufficiently patient, this theory can explain the gradual (rather than immediate) fall in import tariffs. Due to the gradual decline in natural trade frictions, the world has experienced fewer losses in transit, making lower tariffs more self-enforceable. This has allowed successive rounds of trade liberalisation through the GATT/WTO. Second, the theory may be able to explain why the developed world has experienced much deeper reductions in politically induced trade protection, since transport costs are generally much higher between poorer countries. Finally, since the theory suggests that global free trade is more self-enforceable when trade costs are lower, it can help explain the emergence of the large number of FTAs since World War II.

Moreover, I believe the model is relevant for the design of international trading institutions such as the WTO and the European Union. A relevant question is the deepness of political integration and the extent to which institutions such as the EU and the WTO should be given supra-national powers over trade policy. According to the theory, FTAs can be supported for a larger range of discount factors when trade costs are lower. This implies that when trade costs fall the European Union could actually return powers over trade policy to member states, knowing that trade liberalisation stands a higher chance of sustaining itself in the absence of a law enforcement agency.

At the same time, the model can be relevant in times of economic recessions, since the degree to which governments value the future often depends on the state of the macroeconomy. In times of recession, governments may care more about short-run considerations, and for that reason international law enforcement agencies may need to be deepened to deal with falling natural trade costs and the consequent increased incentives to deviate.

The literature on trade policy reveals many channels through which trade costs affect the welfare of import tariffs. In models where there is free entry and exit a standard motive for imposing import tariffs is the so-called home-market effect. Two prominent examples are Venables (1985) and Venables (1987). The first of these models considers an oligopoly model with free entry and homogeneous products while the second uses Dixit-Stiglitz style monopolistic competition. In both models, a positive import tariff imposed by one country has the effect of attracting entry of firms in that country and exit in the other. This increases the amount of goods available not subjected to costly trade costs thus increasing welfare for the country that imposes the tariff. Countries will consequently find themselves caught in a tariff war to attract firms.

Using a model of multinational enterprises, Ludema (2002) constructs a similar model where firms face a trade-off between being close to foreign markets (to avoid trade costs) and to concentrate production at home (to exploit economies of scale). When trade costs are reduced the incentive to unilaterally impose import tariffs decreases since a smaller proportion of the traded quantities is lost in transit and the desire to exploit economies of scale is increased. The desire to cooperate is therefore also greater. Notice that the model in the present paper finds that the incentive to impose import tariffs increases when trade costs fall which is the opposite to Ludema's (2002) finding. It is therefore reassuring that the finding that a FTA agreement is more self-enforceable when trade costs are lower is the same in this model as in Ludema's (2002). But in my model the result has a different driver, namely that the externalities that unilateral trade policy impose internationally when trade costs are lower makes a FTA more self-enforceable.

Ludema's (2002) model produces the counterfactual implication that there is a positive relationship between FDI and import protection. It is well known, however, that countries with larger FDI flows have lower levels protection. Hence, the present paper demonstrates that the relationship between the self-enforceability of FTAs and trade costs carries over to an environment with different motives for trade policy.

More recently, Zissimos (2010) models the formation of FTAs as a coordination problem. He shows how distance can be used to solve such coordination problems between countries. His main finding is that countries that are located closer to each other geographically (distance could here be interpreted as transport costs) impose greater externalities (in the form of terms-of-trade and profit-shifting distortions) upon each other, thus increasing the desire to coordinate their trade policy choices by forming a FTA. As in the present model, however, Zissimos (2010) does not invoke a requirement that trade agreements must be self-enforceable. Other papers which consider similar issues as the present paper in models of perfect competition include Bond and Syropoulos (1996), Bond (2001), and Bond, Syropoulos and Winters (2001).

On the empirical side, very few papers have analysed the relationship between trade costs and cooperative trade policy, but casual observation should convince us that there is indeed a relationship. In fact, of the many variables which might explain the emergence of trade agreements proximity stands out. In the first systematic attempt to model the predictors of FTA membership, Baier and Bergstrand (2004) find that proximity is a good predictor of membership of FTAs. They find this result using a probit model for a sample of 54 countries (or 1431 country pairs).

Since I have restricted this model to a two-country framework, I am not contributing to the debate on regionalism as the absence of a third country omits the possibility of trade diversion. This could potentially affect both the long-run benefits from cooperation and short-run incentives to deviate for the case of preferential trade agreements. If it is assumed that trade agreements are signed between natural trading partners, however, this trade diversion effect may not pose a large problem. In any case, this paper may be more applicable to the case of multilateral trade agreements as negotiated under the GATT/WTO. In the concluding section of this paper I discuss how my results might extend to frameworks with more than two countries.

The present paper also contributes to the literature on gradual trade liberalisation. In Furusawa and Lai (1999), the requirement that trade agreements be self-enforceable induces gradual trade liberalisation. They construct a two-country two-sector trade model in which there is an adjustment cost to be incurred for a worker to move from one sector to the other. If the two countries choose to embark on a path of trade liberalisation they will liberalise as much as possible while keeping each country's incentive to stay in the agreement. In every period, deeper trade liberalisation is made possible and the importable sector shrinks. In a similar model, Maggi and Rodriguez-Clare (2007) show that trade liberalisation can be gradual due to imperfect mobility of capital. This allows faster trade liberalisation when capital is more mobile. That paper, however, does not require trade agreements to be self-enforceable. In the present paper, gradualism emerges from the effects of trade costs: a gradual exogenous deline in trade costs makes lower tariffs more self-enforceable.

The paper is organised as follows. In Section 2 I present the model of oligopoly which will be used throughout the paper. Section 3 discusses the unilateral trade policy equilibrium, before Section 4 moves on to discuss how cooperative trade policy can be used to solve this inefficiency. In Section 5, I discuss how the requirement of self-enforceability can change the outcomes of cooperative trade policy, and Section 6 concludes while discussing some useful extensions.

# 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 consider a world with two countries called Home and Foreign and an infinite number of discrete time periods. Each country has one firm each producing one good. I assume there is no entry/exit such that firms make abnormal profits in equilibrium. Preferences are identical across countries and can be represented by the following quasilinear-quadratic utility function in each period:

$$U\left(\mathbf{q_i}, M_i\right) = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2}\sum_{j=h,f} q_{ij}^2 + M_i, \;\; i = h, f, \tag{1}$$

where $q_{ij}$ is country $i$'s consumption of country $j$'s products, $\mathbf{q_i} \equiv (q_{ih}, q_{if})$ is country $i$'s consumption vector, $Q_i \equiv q_{ih} + q_{if}$ and $M_i$ is country $i$'s consumption of the numeraire good. The numeraire is freely traded across countries to settle the balance of trade, and I assume that each country's endowment of this good is sufficient to guarantee a positive consumption 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.

The two countries may not be symmetric in every aspect but for convenience I shall present all economic expressions for Home as the analogous expressions for Foreign are not hard to express once the reader has been presented with the expressions for Home. By maximising utility in (1) it is possible to derive Home's demand for the Home firm's good and the Foreign firm's good, respectively, as:

$$p_{hh} = a - q_{hh} - \gamma q_{hf} \quad \text{and} \quad p_{hf} = a - q_{hf} - \gamma q_{hh}. \tag{2}$$

The analogous demand functions for Foreign's demand for the Home good and the Foreign good can be found by exchanging $h$ and $f$ in (2). Trade is subject to natural trade costs of the iceberg form. In order for one unit of exports to arrive in Home, $1 + \alpha_h$ units must be produced. Similarly, in order for one unit of exports to arrive in Foreign, $1+\alpha_f$ units must be produced. I assume there are no internal natural trade costs. 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 Home sets a tariff equal to $\tau_h$ on imports from Foreign's firm, and Foreign sets a tariff equal to $\tau_f$. I also assume there are no internal political trade barriers.

Both firms produce at the same marginal cost of $c$ in their respective domestic markets, $c = c_{hh} = c_{ff}$, but due to trade costs (both political and natural) the effective marginal cost of exporting becomes $c_{fh} = c + \alpha_f + \tau_f$ for the Home firm and $c_{hf} = c + \alpha_h + \tau_h$ for the Foreign firm. Markets are segmented and firms compete in a Cournot fashion by choosing quantities in each country. In the Home market, the Home firm solves the following problem, $\max_{q_{hh}} \pi_{hh} = (p_{hh} - c)\, q_{hh}$, and the Foreign firm solves, $\max_{q_{hf}} \pi_{hf} = (p_{hf} - c_{hf})\, q_{hf}$, yielding the following first-order conditions:

$$p_{hh} - c - q_{hh} = 0 \quad \text{and} \quad p_{hf} - c - \alpha_h - \tau_h - q_{hf} = 0\ . \tag{3}$$

Using (2) these conditions can be rewritten as:

$$a - c - 2q_{hh} - \gamma q_{hf} = 0 \quad \text{and} \quad a - c - \tau_h - \alpha_h - 2q_{hf} - \gamma q_{hh} = 0. \tag{4}$$

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

$$q_{hh} = \frac{\Gamma(0,\gamma) + \gamma(\tau_h + \alpha_h)}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} \text{ and } q_{hf} = \frac{\Gamma(0,\gamma) - 2(\tau_h + \alpha_h)}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)}, \tag{5}$$

where $\Gamma(\cdot)$ is defined as $\Gamma(k,\gamma) \equiv 2 - \gamma + k\gamma$, and I have normalised such that $a - c = 1$. By summing the quantities produced by each firm I get an expression for the total quantity demanded in Home:

$$Q_h = \frac{2 - (\tau_h + \alpha_h)}{\Gamma(2,\gamma)}$$

By solving the profit maximising problems for the Home and Foreign firms, respectively, in the Foreign market I obtain the analogous expressions for Cournot quantities in the Foreign market, $q_{ff}$, $q_{fh}$ and $Q_f$.

The equilibrium quantities have the following properties:

$$\begin{aligned}
\frac{dq_{hh}}{d\tau_h} &= \frac{dq_{hh}}{d\alpha_h} = \frac{\gamma}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} > 0 \text{ and } \frac{dq_{hf}}{d\tau_h} = \frac{dq_{hf}}{d\alpha_h} = -\frac{2}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} < 0. \\
\frac{dQ_h}{d\tau_h} &= \frac{dQ_h}{d\alpha_h} = -\frac{1}{\Gamma(2,\gamma)} < 0
\end{aligned}$$

If Home raises its tariff on imports of goods from Foreign, the consumption of foreign imports as well as the total consumption fall, but the consumption of Home's domestic good increases. Exogenous increases in natural trade costs, $\alpha_h$, have the same effect on quantities. In fact, what matters for the equilibrium quantities are *total* trade costs whether political or natural. A similar argument applies to quantities in Foreign.

Using the first-order condition in (3) I obtain the equilibrium per-period profits of the two firms in the Home market:

$$\pi_{hh}(\tau_h, \alpha_h) = (p_{hh} - c)\, q_{hh} = q_{hh}^2 \quad \text{and} \quad \pi_{hf}(\tau_h, \alpha_h) = (p_{hf} - c - \tau_h - \alpha_h)\, q_{hf} = q_{hf}^2 \ .$$

Hence, I have,

$$\frac{d\pi_{hh}}{\tau_{hh}} = \frac{d\pi_{hh}}{\alpha_h} = \frac{2\gamma q_{hh}}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} > 0 \text{ and } \frac{d\pi_{hf}}{\tau_h} = \frac{d\pi_{hf}}{\alpha_h} = -\frac{4q_{hf}}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} < 0. \tag{6}$$

If Home raises its tariff (or there is an exogenous rise in natural trade costs) on imports from Foreign, the Home firm's profits from domestic sales rise, but the export profits of the Foreign firm in Home fall. It is, analogously, possible to express the profits of the Home firm and of the Foreign firm, respectively $\pi_{ff}$ and $\pi_{fh}$, in the Foreign market, and apply the same arguments to profits there.

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.

I define total consumer welfare in Home, $C_h$, to be the consumers surplus enjoyed from consuming every variety (provided $\gamma < 1$) and the tariff revenue which is redistributed back to individuals in a lump-sum fashion. I can express this total consumer welfare in each period as:

$$C_h(\tau_h, \alpha_h) = CS_h(\tau_h, \alpha_h) + TR_h(\tau_h, \alpha_h) = \frac{1}{2}(a - p_{hh})\, q_{hh} + \frac{1}{2}(a - p_{hf})\, q_{hf} + \tau_h q_{hf}. \tag{7}$$

The total profits of Home's firm consist of the profits it makes from serving its domestic market as well as its profits from supplying the market in Foreign. I can express these per-period aggregate profits as:

$$\begin{aligned} \Pi_h(\tau_h, \tau_f, \alpha_h, \alpha_f) &= \pi_{hh}(\tau_h, \alpha_h) + \pi_{fh}(\tau_f, \alpha_f) \qquad (8) \\ &= (p_{hh} - c)\, q_{hh} + (p_{fh} - c - \tau_f - \alpha_f)\, q_{fh}. \end{aligned}$$

The per-period welfare of each country can be expressed by adding up consumer welfare in (7) and profits in (8):

$$W_h(\tau_h, \tau_f, \alpha_h, \alpha_f) = CS_h(\tau_h, \alpha_h) + TR_h(\tau_h, \alpha_h) + \Pi_h(\tau_h, \tau_f, \alpha_h, \alpha_f). \tag{9}$$

Exchanging $h$ and $f$ in equations (8)-(9) gives the corresponding per-period expressions for consumer surplus, tariff revenue and aggregate profits for Foreign.

# 3 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 the two 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 (2011), 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 (9) with respect to $\tau_h$ yields[3],

$$\frac{dW_h}{d\tau_h} = \underbrace{-q_{hf}\frac{dp^*_{hf}}{d\tau_h}}_{ToT\geq 0} + \underbrace{\tau_h\frac{dq_{hf}}{d\tau_h}}_{VoT\leq 0} + \underbrace{(p_{hh} - c)\frac{dq_{hh}}{d\tau_h}}_{PS\geq 0}, \tag{10}$$

where $p^*_{hf}$ is the net-of-tariff price of Foreign's good sold in Home, or $p^*_{hf} = p_{hf} - \tau_h$. The ToT effect is the variation in the net-of-tariff price which Foreign's firm receive for their exports to Home. In this model, the ToT effect is positive such that an increase in Home's import tariff improves Home's terms of trade. The tariff reduces Home'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 $\frac{dq_{hh}}{d\tau_h} = 0$, 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

[3] See Appendix A for the derivation.

(5) and the inverse demand function (2) in (10), I can solve for the optimal non-cooperative tariffs for Home, and analogously for Foreign, as[4]:

$$\tau_h^N = \frac{1-\alpha_h}{3} \text{ and } \tau_f^N = \frac{1-\alpha_f}{3}, \tag{11}$$

where the superscript $N$ on the optimal tariff in (11) which stands for Nash, and is there to illustrate the prisoner's dilemma nature of non-cooperative trade policy. Notice that Home's (Foreign's) optimal tariff is independent of the trade costs incurred by exporting to Foreign (Home). Notice further from Appendix A that the governments do not set Nash tariffs strategically such that the Nash tariff in Home (Foreign) is independent of the Nash tariff in Foreign (Home). This feature of the model is due to the fact that the two countries' markets are segmented. Because of the assumed linear demand function it is possible that tariffs become prohibitive. To rule this out I impose the condition,

$$\alpha_h < 1 - \frac{3}{4}\gamma \left(\alpha_f < 1 - \frac{3}{4}\gamma\right), \tag{12}$$

on Home (Foreign) trade costs throughout the rest of the paper. In Appendix C, I prove the following proposition:

**Proposition 1** *: If $\alpha_h < 1 - \frac{3}{4}\gamma \left(\alpha_f < 1 - \frac{3}{4}\gamma\right)$ there exists a unique non-prohibitive Nash tariff for Home (Foreign).*

It is useful to see how changes in Home trade costs change the incentive to impose import tariffs. Taking the derivative with respect to $\alpha_h$ in (10) yields[5]:

$$\frac{d^2 W_h}{d\tau_h d\alpha_h} = -\frac{1}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} < 0. \tag{13}$$

This implies that as $\alpha_h$ falls, the gain from imposing tariffs 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 terms of trade in Home's favour, and shifting profits towards the domestic firm in the domestic market. Hence, the negative correlation between the Nash tariffs and trade costs, that is $\frac{d\tau_h^N}{d\alpha_h} = -\frac{1}{3} < 0$ (and $\frac{d\tau_f^N}{d\alpha_f} = -\frac{1}{3} < 0$).

Unilateral trade policy is inefficient, however, as one country's welfare gain comes at the expense of the other. By taking the derivative of Home's welfare with respect to the Foreign

[4]See Appendix B for the derivation.

[5]See Appendix B for the derivation.

import tariff 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[6]:

$$\frac{dW_h}{d\tau_f} = \underbrace{q_{fh}\frac{dp^*_{fh}}{d\tau_f}}_{ToT\leq 0} - \underbrace{(\tau_f + \alpha_f)\frac{dq_{fh}}{d\tau_f}}_{VoT\geq 0} + \underbrace{(p_{fh} - c)\frac{dq_{fh}}{d\tau_f}}_{PS\leq 0}, \tag{14}$$

where $p^*_{fh}$ is the net-of-tariff price of Home's good sold in Foreign. Substituting Foreign's Nash tariff in (11) into (14) yields[7]:

$$\left.\frac{dW_h}{d\tau_f}\right|_{\tau_f^N} = \frac{4\left[\gamma - 4\left(1-\alpha_f\right)\right]}{3\left[\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right]^2} < 0 \tag{15}$$

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

$$\left.\frac{d^2W_h}{d\tau_f d\alpha_f}\right|_{\tau_f^N} = \frac{16}{3\left[\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right]^2} > 0.$$

It is similarly possible to see how the substitution index affects international externalities by taking the derivative of (15) with respect to $\gamma$:

$$\left.\frac{d^2W_h}{d\tau_f d\gamma}\right|_{\tau_f^N} = \frac{16\left(1-\gamma^2\right) + 64\gamma\left(1-\alpha_f\right) - 4\gamma^2}{3\left(\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right)^3} > 0.$$

An increase in the degree of strategic interaction among firms, thus, increases international externalities. A trade agreement may be used to overcome this prisoner's dilemma by inducing cooperation.

# 4 Trade liberalisation with commitment

In this section, I derive the most efficient tariffs which would obtain if the two countries were able to commit themselves to all future tariffs as stipultated by a trade agreement. This implies that Home and Foreign are not allowed to deviate from an agreed-upon tariff. This case would arise, for example, if there was a supra-national government, or some international law enforcement agency which was able to ensure that agreements are honoured. Given this environment, the governments of Home and Foreign set two cooperative tariffs, $\tau_h^C$ and $\tau_f^C$, to maximise the present discounted value of their joint welfare:

$$\max_{\tau_h,\tau_f}\frac{1}{1-\delta}J\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right) = \frac{1}{1-\delta}\left(W_h\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right) + W_f\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right)\right), \tag{16}$$

[6] See Appendix A for the derivation.
[7] See Appendix B for the derivation.

where $\delta < 1$ is the discount factor assumed symmetric across countries. As with unilateral policy the governments move first by setting efficient tariffs and firms then choose quanities subject to those tariffs.Taking the derivative of (16) with respect to $\tau_h$ yields the following first order condition[8]:

$$\frac{dJ}{d\tau_h} = (p_{hh} - c)\frac{dq_{hh}}{d\tau_h} + (p_{hf} - c)\frac{dq_{hf}}{d\tau_h} - \alpha_h \frac{dq_{hf}}{d\tau_h}. \tag{17}$$

Exchanging $h$ and $f$ in (17) gives the analogous first order condition for the Foreign cooperative tariff. Substituting the inverse demand functions (2) and the Cournot quantities (5) into (17), I can solve for the internationally efficient tariff for Home, and analogously for Foreign, as[9]:

$$\tau_h^E = \tau_f^E = \frac{-\Gamma(0,\gamma)\Gamma(0,\gamma) + \frac{1}{2}\alpha_h(4+\gamma^2) + \frac{1}{2}\alpha_f(4+\gamma^2)}{(4-3\gamma^2)}. \tag{18}$$

Thus, if Home and Foreign were able to commit to any tariff they would choose $\tau_h^E$ and $\tau_f^E$, respectively, in every period forever. In the absence of any trade costs $\alpha_h = \alpha_f = 0$, the efficient cooperative tariffs are negative. This is because in the presence of oligopolistic markets, firms produce suboptimal quantities and it is therefore efficient to subsidise them. As trade costs increase (both Home and Foreign, respectively, $\alpha_h$ and $\alpha_f$), however, an increasing amount of the traded quantities are lost in transit. Hence, from a joint welfare-maximising perspective it becomes optimal to subsidise the traded quantities less, and as trade costs reach a critical threshold it will be more efficient to impose a positive tariff to reduce the quantity lost in transit. If trade costs are symmetric, this critical threshold can be solved from (18) as,

$$\alpha \geq \overline{\alpha} \equiv 1 - \frac{4\gamma}{4+\gamma^2}. \tag{19}$$

In Appendix D, I show that as trade costs symmetrically approach their upper bounds in (12), the efficient tariffs in (18) converge to the Nash tariffs in (11), and in the limit they are equal:

$$\underset{\alpha\to 1-\frac{3}{4}\gamma}{\tau_h^E} = \tau_h^N \text{ and } \underset{\alpha\to 1-\frac{3}{4}\gamma}{\tau_f^E} = \tau_f^N \tag{20}$$

Notice the conflict of interest between non-cooperative and cooperative trade policy. When acting unilaterally countries wish to set higher tariffs when trade costs fall, but from a bilateral perspective tariffs should be lowered when trade costs fall. In order to get a feel

[8] See Appendix D for the derivation

[9] See Appendix D for the derivation.

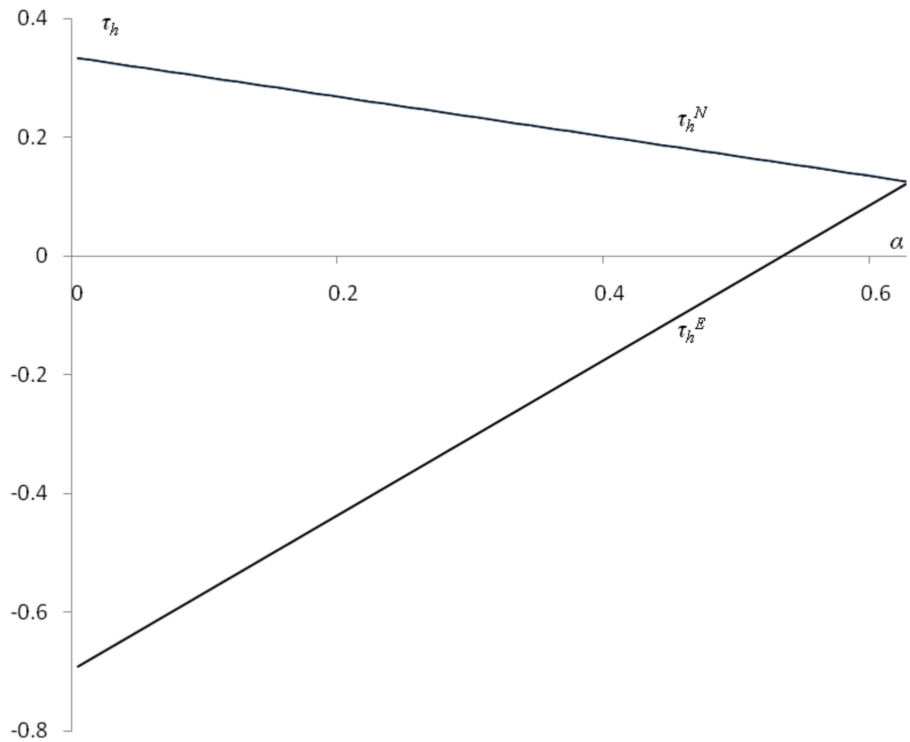


Figure 2: The relationship between tariffs and trade costs. ($\gamma = 0.5$).

for the differences of unilateral and cooperative trade policy objectives, I draw the Nash and efficient tariffs as functions of trade costs in Figure 2. I have set $\gamma = 0.5$ such that trade is eliminated when $\alpha = 1 - \frac{3}{4}\frac{1}{2} = \frac{5}{8}$ (see Eq. (12)). The reason for setting $\gamma$ equal to this value is purely for illustrative purposes, and it could be set at any other value yielding similar results.

# 5 Trade liberalisation without commitment

In this paper, focus lies primarily on cooperative tariffs which can be sustained through a reputational mechanism. This implies that Home and Foreign do not necessarily achieve the highest possible level of long-run welfare in any agreement as in the previous section. This is because in every period, as will be clear below, each country is weighting the short-run gain from deviating from an agreed-upon tariff against the long-run gain from adhering to it. Thus, because of the requirement of self-enforceability, one could say that from a long-run prespective, the two countries may be constraining themselves to a second-best agreement.

In this section, I consider two types of agreements which are subject to a self-enforcement constraint. First, I solve for the optimal bilateral tariff which can be sustained through a reputational mechanism. This optimum tariff depends crucially on the degree of patience of the two countries. Patient countries are able to agree upon tariffs which are closer to the efficient tariffs of the previous section, whereas impatient countries cannot sustain significant departures from the Nash tariffs. Second, I will analyse the self-enforceability of global free trade. This particular case is useful since many countries have signed FTAs in the post-war era. Since a FTA involves removing import tariffs as a trade policy instrument, it could be argued that such an agreement is the most self-enforceable if there is a fixed cost associated with reinstating customs and border controls. I am not modelling this fixed cost explicitly but it could easily be added to the analysis. In both cases, I will examine how trade costs affect the incentives to engage in cooperative trade policy.

As with the case of unilateral trade policy, the setting of cooperative trade policy can be considered as a two-stage game: in the first stage governments jointly set tariffs, and in the second the two firms choose Cournot output levels in, respectively, the Home and Foreign market. The two countries play this two-stage game in every period an inifinite number of times. A strategy is an infinite sequence of functions mapping the history of play into current actions. Let the cooperative tariffs be given as the pair $\left(\tau_h^C, \tau_f^C\right)$, such that the present discounted value of adhering to the agreement is:

$$\frac{1}{1-\delta} W_h^C \left(\tau_h^C, \tau_f^C, \alpha_h, \alpha_f\right). \tag{21}$$

I assume that if a country does not honour the trade agreement in any given period, the other country will punish it by reverting to the politically optimal Nash tariff forever as of the following period. The deviating country will, however, enjoy a short run benefit from deviation by distorting the terms-of-trade in its favour and shifting profits towards the domestic industry. It is possible to find this one-shot deviation payoff from the point of view of Home by substituting Home's non-cooperative Nash tariff from (11) into the expression for welfare:

$$W_h^D \left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right),$$

where the superscript $D$ stands for deviation. Since the other country will punish the cheating nation by reverting to the Nash tariff forever as of the following period, the deviation payoff reduces to the following after one period:

$$W_h^P \left(\tau_h^N, \tau_f^N, \alpha_h, \alpha_f\right),$$

where the superscript $P$ stands for punishment. The present discounted value of the deviation welfare in Home can thus be expressed as:

$$W_h^D\left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right) + \frac{\delta}{1-\delta} W_h^P\left(\tau_h^N, \tau_f^N, \alpha_h, \alpha_f\right). \tag{22}$$

In order for a trade agreement to be self-enforceable, the welfare from honouring the trade agreement must be greater than or equal to the welfare from deviation. The discounted value of the welfare from cooperation can be obtained from (21) and together with (22) I can express the following self-enforcement constraint for Home:

$$\frac{1}{1-\delta} W_h^C\left(\tau_h^C, \tau_f^C, \alpha_h, \alpha_f\right) \geq W_h^D\left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right) + \frac{\delta}{1-\delta} W_h^P\left(\tau_h^N, \tau_f^N, \alpha_h, \alpha_f\right). \tag{23}$$

## 5.1 Optimal self-enforceable tariffs: explaining gradualism in trade liberalisation

The efficient tariffs obtained in Section 4 may not be supported in an agreement which is subject to a self-enforcement constraint for all discount factors. It may be that higher tariffs can be agreed-upon, however, which can be sustained by repeated interaction. In this subsection, I define an optimal self-enforceable tariff and examine the relationship between this tariff and natural trade costs. For the purposes of this subsection I assume that trade costs are symmetric such that $\alpha_h = \alpha_f = \alpha$. Before proceeding, however, it will be useful to calculate the level of patience required to sustain the internationally efficient tariffs in (18). This can be done by solving (23) for a critical discount factor, $\delta_e$, above which an agreement stipulating that Home and Foreign adhere to the efficient tariffs, $\tau_h^E$ and $\tau_f^E$, respectively, is self-enforceable:

$$\delta_e = \frac{W_h^D\left(\tau_h^N, \tau_f^E, \alpha\right) - W_h^C\left(\tau_h^E, \tau_f^E, \alpha\right)}{W_h^D\left(\tau_h^N, \tau_f^E, \alpha\right) - W_h^P\left(\tau_h^N, \tau_f^N, \alpha\right)}. \tag{24}$$

Notice that the numerator of (24), $W_h^D - W_h^C$, is the short-run benefit to Home from deviating from the internationally efficient tariff when Foreign cooperates. This is given as the short run benefit from switching the terms of trade in Home's favour, and shifting profits towards the domestic firm in the Home market relative to the efficient tariffs. The denominator of (24), $W_h^D - W_h^P$, is the long-run loss in welfare when Foreign, which was caught by surprise in the previous period, retaliates by reverting to the Nash tariff. The critical discount factor is thus given as the ratio of the short-run gain from deviation to the long-run loss. In Appendix E, I demonstrate that this critical discount factor is,

$$\delta_e = \frac{3\left(4-\gamma^2\right)}{2\left(8-3\gamma^2\right)}. \tag{25}$$

Hence, for any discount factor greater than $\delta_e$ the internationally efficient tariffs can be sustained. Notice that this discount factor is independent of $\alpha$. When trade costs fall, the temptation to cheat on the agreement increases but so does the long-run benefit of adhering to it, leaving the critical discount factor unchanged. Notice further that $\delta_e$ is increasing in $\gamma$[10]. When $\gamma$ is larger, the temptation to unilaterally deviate from an agreement by distorting terms of trade and shifting profits towards the domestic firm is larger raising the level of patience required to sustain the efficient tariffs.

If the two countries discount the future at a level lower than $\delta_e$, however, it is then possible to find a tariff which improves welfare whilst keeping each country's incentive to stay in the agreement? Let the optimal self-enforceable tariffs for Home and Foreign be given as the pair $\left(\tau_h^S, \tau_f^S\right)$. To fix ideas, I rewrite the self-enforcement constraint (23) for Home in the following convenient way:

$$W_h^D\left(\tau_h^N, \tau_f^S, \alpha\right) - W_h^C\left(\tau_h^S, \tau_f^S, \alpha\right) \leq \delta\left(W_h^D\left(\tau_h^N, \tau_f^S, \alpha\right) - W_h^P\left(\tau_h^N, \tau_f^N, \alpha\right)\right), \tag{26}$$

and symmetrically I can write the self-enforcement constraint for Foreign by exchanging $h$ and $f$ as:

$$W_f^D\left(\tau_f^N, \tau_h^S, \alpha\right) - W_f^C\left(\tau_f^S, \tau_h^S, \alpha\right) \leq \delta\left(W_f^D\left(\tau_f^N, \tau_h^S, \alpha\right) - W_f^P\left(\tau_f^N, \tau_h^N, \alpha\right)\right), \tag{27}$$

Home and Foreign would jointly like to achieve the highest level of long-run welfare by maximising joint welfare (16). The requirement of self-enforceability, however, constrains the two countries by (26) and (27). The optimal self-enforceable tariff solves the following:

$$\begin{aligned} \max_{\tau_h, \tau_f} \frac{1}{1-\delta} J\left(\tau_h, \tau_f, \alpha\right) &= \frac{1}{1-\delta}\left(W_h\left(\tau_h, \tau_f, \alpha\right) + W_f\left(\tau_h, \tau_f, \alpha\right)\right) \\ & s.t.\ (26) \text{ and } (27). \end{aligned} \tag{28}$$

If the discount factor $\delta \geq \delta_e$, the constraints in (26) and (27) are not binding, and Home and Foreign will set import tariffs at their efficient levels given in (18). Otherwise they will set import tariffs at their minimum self-enforceable levels. In Appendix E, I show that the minimum self-enforceable tariffs which solve (28) when (26) and (27) are binding, are given as:

$$\tau_h^M = \tau_f^M = \frac{\alpha\left(40\delta + 3\gamma^2 - 12\right) - 40\delta - 3\gamma^2 + 12 + 24\delta\gamma}{3\left(12 - 3\gamma^2 - 8\delta\right)}. \tag{29}$$

Thus, the optimal self-enforceable tariffs for Home and Foreign, respectively, are:

$$\tau_h^S = \max\left(\tau_h^M, \tau_h^E\right) \text{ and } \tau_f^S = \max\left(\tau_f^M, \tau_f^E\right)$$

[10] The derivative $\frac{d\delta_e}{d\gamma} = \frac{12\gamma}{(8-3\gamma^2)^2} \geq 0$ for $\gamma \in [0; 1]$.

In order to get a feel for these tariffs, imagine two extreme worlds. The first world is characterised by extreme myopia, such that Home and Foreign care only about welfare in the current period ($\delta = 0$). Setting $\delta = 0$ in (29) it is clear that in this case, the self-enforceable tariffs are equal to the Nash tariffs. If the world cares only about current-period welfare, no departures from the Nash tariffs are possible. In the second world the two countries are characterised by farsightedness with a discount factor in the range, $\delta \in ]\delta_e; 1[$. In this case, the two countries are patient enough to implement the efficient tariffs in (18). Thus, in an extremely myopic world, no agreement is possible since only the Nash tariffs are sustainable, but in a very farsighted world, the efficient tariffs can be implemented. But how about discount factors that lie in between these cases? Assume now that the world discounts the future at the rate $\delta \in ]0; \delta_e[$. This is where the minimum cooperative tariffs (29) become important. These tariffs can be thought of as a weighted average of the Nash tariffs and the efficient tariffs. A property of the minimum tariffs which I prove in Appendix E is that, when trade costs increase towards the upper bound they converge towards the Nash tariffs:

$$\underset{\alpha \to 1-\frac{3}{4}\gamma}{\tau_h^M} = \tau_h^N \text{ and } \underset{\alpha \to 1-\frac{3}{4}\gamma}{\tau_f^M} = \tau_f^N. \tag{30}$$

This is due to the property that there is no conflict between unilateral and cooperative trade policy in the limit where import tariffs are prohibitive. Formally, I can find out the relationship between trade costs and trade policy by taking the derivative of (29) wrt. $\alpha$. This yields:

$$\frac{d\tau_h^M}{d\alpha} = \frac{(40\delta + 3\gamma^2 - 12)}{3\,(12 - 3\gamma^2 - 8\delta)}.$$

Since the denominator of this derivative is positive, I can evaluate the sign by solving for $\delta$ in the numerator. Hence, there is a positive relationship between $\tau_h^M$ (and symmetrically for $\tau_f^M$) and $\alpha$ if

$$\delta > \frac{12 - 3\gamma^2}{40}. \tag{31}$$

For convenience, I am going to define medium to high discount factors as any discount factor greater than the threshold given in (31). Any discount factor lower than or equal this threshold is defined as a low discount factor. With a slight abuse of terminology, I am going to say that the world is patient if they have medium to high discount factors, and impatient if they have low discount factors. Hence, if the countries are patient, and if they implement the tariffs, $\tau_h^S$ and $\tau_f^S$, respectively, they will set lower tariffs when trade costs fall. If, on the other hand, countries are impatient, they may still find cooperative tariffs which can make them both better off. These cooperative tariffs, however, will need to be raised in response to falling trade costs. The reason for these relationships are as follows. As can be seen from (26) the greater is the discount factor, the greater the valuation the two countries will place on their welfare in the long run. This implies that for a patient world, the cooperative tariff must be closer to the efficient tariffs in (18), which are increasing in trade costs. In this patient world, therefore, the objective of Home and Foreign is mainly to use import tariffs jointly to minimise the transit losses associated with trade costs. In an impatient world, on the other hand, the cooperative tariffs must be closer to the Nash tariffs, which are decreasing in trade costs. In this case, if tariffs depart significantly from their Nash levels, the temptation to distort terms of trade and to shift profits towards the domestic firms is too large. Hence, in a patient world, trade liberalisation is deeper when natural trade costs fall. Provided WTO member countries have been sufficiently patient, this may help explaining gradualism in the trade liberalisation rounds under the GATT/WTO. Falling trade costs, as evidenced in Figure 1, may have facilitated deeper self-sustainable trade liberalisation over time. Impatient countries may also achieve trade liberalisation according to this theory, but stronger international law enforcement agencies are required to sustain welfare-enhancing trade liberalisation in response to falling trade costs.

In Figure 3 I depict the optimal self-enforceable tariff in Home as a function of trade costs for various discount factors in the range, $\delta \in ]0; \delta_e[$. The two bold lines are the efficient and Nash tariffs, respectively, reproduced from Figure 2. The two dashed lines represent optimal self-enforceable tariffs for discount factors of 0.2 and 0.5, respectively. When the discount factor is 0.2, short-run considerations are more important. In fact, in order to keep each country's incentive to stay in the agreement, tariffs need to be raised in response to falling trade costs. In other words, because the future is not important enough, tariffs must be raised in order to make sure each country does not fall for the temptation to distort the terms of trade in their favour, or shift profits towards their respective domestic industries. On the other hand, when the world discounts the future at the rate 0.5, gradual trade liberalisation is made possible by falling natural trade costs.

## 5.2 The self-enforceability of a FTA

In this subsection I will consider the sustainability of global free trade. This case is useful for at least two reasons: (a) many countries have signed FTAs in the post-war era, and (b)

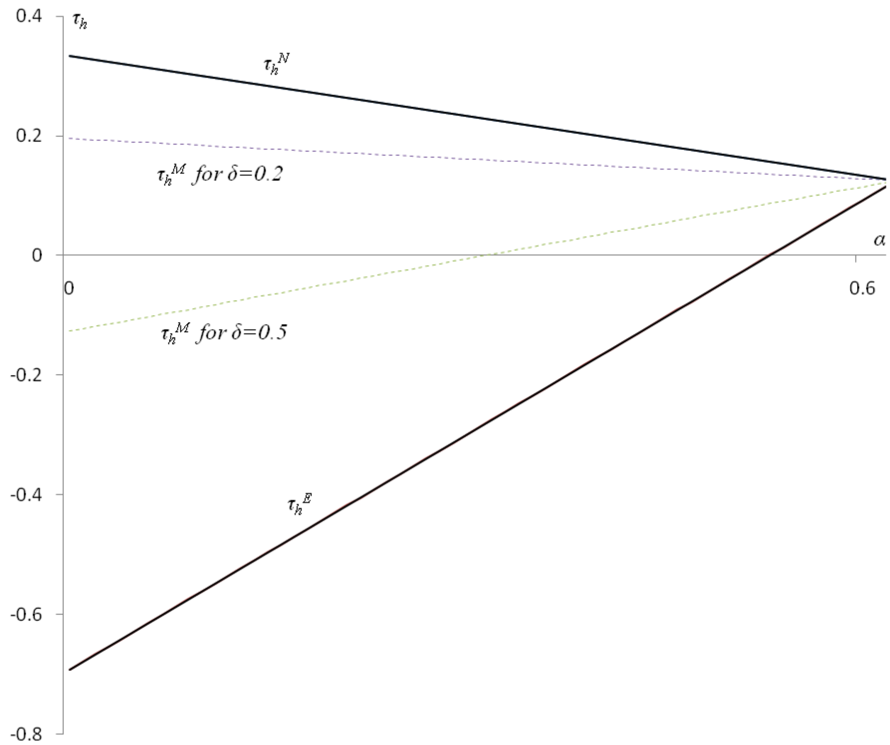


Figure 3: Optimal self-enforcable tariffs. ($\gamma = 0.5$).

a FTA involves giving up import tariffs as a policy instrument as opposed to the efficient tariffs which involve using import tariffs to reach the maximum level of joint welfare. If customs and border controls involve significant fixed costs, it may be that a FTA is more self-enforceable than any other agreement by removing import tariffs as a policy option. Global free trade, however, does not solve the problem of suboptimal quantities which is present under oligopolistic competition. Recall from Section 4 that the efficient tariffs in (18) may be positive for sufficiently high trade costs since the two countries jointly try to minimise traded quantities lost in transit. Hence, it will be the case that from a joint welfare-maximising perspective, free trade is far too deep a trade liberalisation once trade costs reach that threshold (see Eq. (19)). It is, therefore, questionable whether the two countries would choose to sign a FTA at all in this situation. I am going to argue that the two countries would consider signing a FTA so long as the efficient tariffs are non-positive.

The cooperative policy options available to Home and Foreign are, respectively, the tariffs $\tau_h^{FT} = \tau_f^{FT} = 0$ provided $\tau_h^E \leq 0$ and $\tau_f^E \leq 0$. It will be convenient to express the self-enforcement constraint (23) in terms of a critical discount factor, $\delta_c$, above which the FTA is self-enforceable:

$$\begin{aligned}\delta_c(\alpha_h,\alpha_f) &= \frac{W_h^D\left(\tau_h^N,\tau_f^{FT},\alpha_h,\alpha_f\right)-W_h^C\left(\tau_h^{FT},\tau_f^{FT},\alpha_h,\alpha_f\right)}{W_h^D\left(\tau_h^N,\tau_f^{FT},\alpha_h,\alpha_f\right)-W_h^P\left(\tau_h^N,\tau_f^N,\alpha_h,\alpha_f\right)} \\ &= \frac{CS_h\left(\tau_h^N,\alpha_h\right)-CS_h\left(\tau_h^{FT},\alpha_h\right)+TR_h\left(\tau_h^N,\alpha_h\right)+\pi_{hh}\left(\tau_h^N,\alpha_h\right)-\pi_{hh}\left(\tau_h^{FT},\alpha_h\right)}{\pi_{fh}\left(\tau_f^{FT},\alpha_f\right)-\pi_{fh}\left(\tau_f^N,\alpha_f\right)}.\end{aligned} \tag{32}$$

As with (24) the numerator of (32), $W_h^D - W_h^C$, is the short-run benefit from deviating from the FTA when the other country cooperates. This is given as the short run benefit from switching the terms of trade in Home's favour, and shifting profits towards the domestic firm in the Home market relative to free trade. This gain depends only on trade costs incurred by exporting to Home, $\alpha_h$, and it is independent of $\alpha_f$. The denominator of (32), $W_h^D - W_h^P$, is the long-run loss in welfare when the other country, which was caught by surprise in the previous period, retaliates by reverting to the Nash tariff. This loss depends on natural trade costs incurred by exporting to Foreign, $\alpha_f$, but is independent of $\alpha_h$. The critical discount factor is thus given as the ratio of the short-run gain from deviation to the long-run loss.

The next step in the analysis is to examine how trade costs affect the critical discount factor (32) for zero cooperative tariffs, and I will first consider asymmetric changes in trade costs. Suppose there is an exogenous change in trade costs incurred by exporting to Home, $\alpha_h$, possibly due to changes in technology. The short-run gain from deviating from free trade, $W_h^D - W_h^C$, consists of a consumer and a producer gain. The total benefit to consumers from import protection relative to free trade, $CS_h\left(\tau_h^N,\alpha_h\right)-CS_h\left(\tau_h^C,\alpha_h\right)+TR_h\left(\tau_h^N,\alpha_h\right)$, depends on how Home's natural trade costs affect the terms of trade gain and the volume of trade loss relative to free trade. Similarly, the producer gain, $\pi_{hh}\left(\tau_h^N,\alpha_h\right)-\pi_{hh}\left(\tau_h^C,\alpha_h\right)$, depends on how Home's natural trade costs affect the degree to which the Nash tariff is able to shift profits towards the domestic firm relative to free trade. This producer gain is positive provided that firms interact strategically ($\gamma > 0$). In Appendix F I show that there is a negative correlation between Home trade costs, $\alpha_h$, and the short-run benefit from deviation, $W_h^D - W_h^C$. This implies that when $\alpha_h$ is low, the gain from unilaterally deviating from free trade is higher implying that, *ceteris paribus*, a FTA becomes less self-enforceable when Home trade costs are lower. 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 terms of trade in Home's favour, and shifting profits towards the domestic firm in the Home market relative to free trade. This feature of the model is different from Ludema's (2002) model of trade with multinational firms in which the short run gain from deviation is *increasing* in trade costs. The reason for this is that more multinationals increases the variety of goods not subjected to trade costs, such that the gain to consumers is lower when trade costs are lower.

Suppose, now, there is an exogenous change in the natural trade costs incurred by exporting to Foreign, $\alpha_f$. This affects the long-run loss in Home welfare when Foreign retaliates, $W_h^{SR} - W_h^P$. This loss consists of the difference between the Home firm's export profits under free trade trade and under protectionism, $\pi_{fh}\left(\tau_f^C, \alpha_f\right) - \pi_{fh}\left(\tau_f^N, \alpha_f\right)$. In Appendix F, I show that there is a negative correlation between Foreign trade costs and the long-run loss from deviation. The intuition for this result is the following: if trade costs incurred by exporting to Foreign are lower, profits are to a lesser extent naturally diverted away from the Home firm in the Foreign market, thus ensuring the Home firm's profits in the Foreign market are larger. This, however, makes retaliatory responses by Foreign larger. Thus, when Foreign natural trade costs decline, a free trade agreement becomes more self-enforceable by lowering the critical discount factor, $\delta_c$. I summarise my findings in the following proposition:

**Proposition 2** *If trade costs incurred by exporting to Home decrease, a FTA becomes less self-enforceable since it raises the critical discount factor,* $\frac{d\delta_c}{d\alpha_h} < 0$. *If trade costs incurred by exporting to Foreign decrease, however, a FTA becomes more self-enforceable,* $\frac{d\delta_c}{d\alpha_f} > 0$.

The proof of this proposition can be found in Appendix F.

It would be realistic to assume, however, that in the real world, there is not so much asymmetry in terms of trade costs, in particular between developed nations. In the following, therefore, an assumption of symmetry is imposed, that is, $\alpha_h = \alpha_f = \alpha$, and equiproportional changes in trade costs are examined. Notice that both the numerator and the denominator of (32) are decreasing in $\alpha$. Hence, when $\alpha$ decreases the short run gain from deviating from free trade will increase as will the long run punishment. Hence, there are two forces in play: (a) when there is fall in $\alpha$, the short-run gain from deviating from free trade increases, making the free trade agreement *less* self-enforceable, and (b) when there is a fall in $\alpha$, the long-run loss from deviating will decrease making the free trade agreement *more* self-enforceable. Thus, I have to examine how the gain (numerator) and the loss (denominator) vary proportionally with trade costs.

To find the net effect I write up an expression for the critical discount factor[11]:

$$\begin{aligned}
\delta_c(\alpha) &= \frac{CS_h\left(\tau_h^N,\alpha\right) - CS_h\left(\tau_h^{FT},\alpha\right) + TR_h\left(\tau_h^N,\alpha\right) + \pi_{hh}\left(\tau_h^N,\alpha\right) - \pi_{hh}\left(\tau_h^C,\alpha\right)}{\pi_{fh}\left(\tau_f^{FT},\alpha\right) - \pi_{fh}\left(\tau_f^N,\alpha\right)} \\
&= \frac{3(1-\alpha)(4-\gamma^2)}{8(5(1-\alpha)-3\gamma)}. \qquad (33)
\end{aligned}$$

Taking the derivative with respect to $\alpha$ yields:

$$\frac{d\delta_c}{d\alpha} = \frac{30\gamma(4-\gamma^2)}{8(5(1-\alpha)-3\gamma)^2} \geq 0. \qquad (34)$$

On the basis of (34), I can propose the following:

**Proposition 3** *Provided $\gamma > 0$ a FTA becomes more self-enforceable when trade costs between two countries decline, $\frac{d\delta_c(\alpha)}{d\alpha} > 0$. If $\gamma = 0$ there is no relationship between the self-enforceability of a FTA and trade costs.*

A proof of this proposition can be found in Appendix F. Provided there is strategic interaction among firms it will be the case that when trade costs decline, tariffs are more effective at distorting the terms of trade in Home's favour and at shifting profits towards the domestic firm. The long run consequences of unilateral deviation are more severe when trade costs are lower, however, making the long-run benefit from trade cooperation higher.

If there is no strategic interaction between firms ($\gamma = 0$), neither political trade costs nor natural trade costs are able to shift profits towards the domestic firm (see (6)), and it is possible to rewrite the critical discount factor (33) as:

$$\delta_c(\alpha) = \frac{CS_h\left(\tau_h^N,\alpha\right) - CS_h\left(\tau_h^C,\alpha\right) + TR_h\left(\tau_h^N,\alpha\right)}{\pi_{fh}\left(\tau_f^C,\alpha\right) - \pi_{fh}\left(\tau_f^N,\alpha\right)} = \frac{3}{10}. \qquad (35)$$

When $\gamma = 0$, the effect of $\alpha$ on the gain to consumers from deviating is exactly proportional to the loss in export profits. This is because when goods are independent of each other, the change in welfare from natural trade costs affects only traded goods. Thus, the change in the gain to consumers from deviating is proportional to the change in the loss inflicted upon the Foreign firm's export profits, and by symmetry this loss is equal to the Home firm's loss in the Foreign market. The critical discount factor is therefore independent of $\alpha$ and equal to a constant, in this model $\frac{3}{10}$.

When $\gamma$ increases the degree of strategic interaction between firms increases, and profit-shifting becomes a more important motive in the setting of trade policy. Therefore, a trade

[11] See Appendix F for the steps behind this.

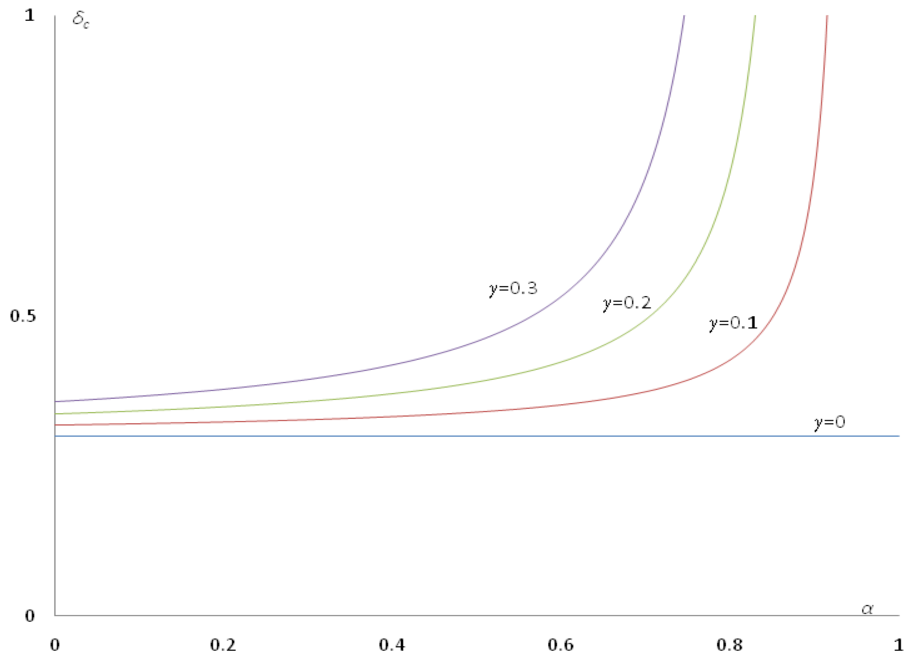


Figure 4: The critical discount factors as functions of trade costs.

agreement becomes less self-enforceable for every value of $\alpha$. This can be seen by taking the derivative of (33) with respect to $\gamma$:

$$\frac{d\delta_c}{d\gamma} = \frac{3\left(1-\alpha\right)\left(12+3\gamma^2-10\gamma\left(1-\alpha\right)\right)}{8\left(5\left(1-\alpha\right)-3\gamma\right)^2} > 0$$

It can also be seen from (34) that the relationship between $\delta_c$ and $\alpha$ becomes steeper when $\gamma$ is higher. Taking the derivative of (34) with respect to $\gamma$ yields,

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

In order to get a feel for the relationship between $\delta_c$ and $\alpha$ I draw them as a function of each other in Figure 4 for different values of $\gamma$. From the diagram it is clear that for $\gamma = 0$ the relationship is a flat line whereas when $\gamma$ increases the relationship between $\delta_c$ and $\alpha$ becomes steeper. When the degree of strategic interaction between firms is higher, the profit-shifting motive for trade policy becomes greater, raising the critical discount factor for every value of $\alpha$. However, a higher $\gamma$ will raise the critical discount factor less for lower trade costs. This is because the problem of profit-shifting is larger when trade costs are lower increasing the incentive to cooperate. This implies that, in this model, FTAs have a greater regional bias when the degree of strategic interaction among firms is higher.

It is possible that trade costs reach a threshold, $\widehat{\alpha}$, above which a FTA is not self-enforceable at all. This threshold can be obtained from (33):

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

Solving for $\alpha$ yields:

$$\alpha \geq \widehat{\alpha} \equiv 1 - \frac{24\gamma}{28+3\gamma^2}.$$

When $\alpha$ is above this threshold, global free trade is too deep a trade liberalisation. From a joint welfare-maximising perspective, too many units of production are lost in transit at such high trade costs, and it is therefore more optimal to impose a positive import tariff closer to the Nash tariffs.

# 6 Concluding remarks

In this paper I have analysed the nature of the relationship between trade costs and trade policy under oligopoly. I first showed the difference of objectives of uniltareal and cooperative trade policy. First, unilateral trade policy maximises each country's domestic welfare with respect to an import tariff. The import tariff distorts the terms of trade and shifts profits towards each country's respective domestic firms. This import tariff is more effective at accomplishing these aims when the degree of natural distortion through trade costs is lower. Hence, each country would like to set higher tariffs when trade costs are lower. Second, when the two countries set import tariffs cooperatively these externalities are neutralised, and the resulting internationally efficient tariffs are lower when trade costs fall since their objective is to minimise losses in transit. Hence, the objectives of cooperative and non-cooperative trade policy diverge when trade costs fall: acting unilaterally, the two countries would like to raise tariffs in response to falling trade costs, whereas the reverse is the case cooperatively. I then added a requirement that trade agreements be sustained under a reputational mechanism. I assumed that if either of the two countries defected from an agreement in any given period, the other would punish it by reverting to the politically optimal Nash tariff forever as of the following period. I analysed two types of such agreements. First, I considered the optimal self-enforceable tariff and demonstrated that provided the two countries care enough about the future, lower import tariffs can be supported when trade costs fall. If the world was too impatient, however, tariffs needed to be raised in response to falling trade costs to keep each country's incentive to remain in the agreement. This is because impatient countries would find it harder to resist the increased short runs gains from deviation when trade costs fall. Second, I considered a FTA, and looked at how changes in trade costs affected the sustainability of such an agreement. When trade costs fall, the one-shot benefit from deviating from a FTA increased, but so did the long run benefit from cooperation. This is because the international externalities which unilateral deviation imposes lowers the incentives to deviate.

The model could be extended in several interesting ways. First, I have ignored potential and interesting asymmetries between countries. Imagine, for example, that each country hosts more than one firm, and that the number of firms is larger in one country than to the other. In this case the country with fewer firms will suffer less when the other country retaliates by reducing market access. It may then be that the small country will be less likely to honour a FTA than the larger country. Second, the model could be extended to include several countries, who sign up for mutually beneficial tariff reductions. In this case, consider one home country and $n-1$ other countries. I could allow trade costs between Home and each of the $n-1$ countries to differ. Imagine that all of the countries come together in a GATT/WTO framework to sign a FTA. Provided that trade costs between Home and each of the $n-1$ members are not so high that a FTA is not optimal from a joint welfare-maximising perspective, such a scenario would yield the some results as in the present two country framework. If one country is so remote (in terms of trade costs) that free trade is not optimal, however, higher tariffs would need to be negotiated with this country in order to ensure sustainability. Third, consider a three country version of this paper, where two of the three countries decide to sign a trade agreement without the third one. Such an agreement would not necessarily be welfare improving if substantial trade is diverted away from the third country. This would have consequences for both the short-run incentives to deviate and the long run benefits from deviation. It would be interesting to see how the results of the present paper would extend to such a setting.

**Derivation of Eq. (10) and Eq. (14)**

Substituting (7) and (8) into (9) yields:

$$\begin{aligned} W_h &= \frac{1}{2}(a-p_{hh})\,q_{hh}+\frac{1}{2}(a-p_{hf})\,q_{hf}+\tau_h q_{hf} \\ &\quad+(p_{hh}-c)\,q_{hh}+(p_{fh}-c-\tau_f-\alpha_f)\,q_{fh}. \end{aligned} \tag{36}$$

Taking the derivative wrt. $\tau_h$ I obtain:

$$\begin{aligned} \frac{dW_h}{d\tau_h} &= -\frac{1}{2}\frac{dp_{hh}}{d\tau_h}q_{hh}+\frac{1}{2}(a-p_{hh})\frac{dq_{hh}}{d\tau_h}-\frac{1}{2}\frac{dp_{hf}}{d\tau_h}q_{hf}+\frac{1}{2}(a-p_{hf})\frac{dq_{hf}}{d\tau_h} \\ &\quad+q_{hf}+\tau_h\frac{dq_{hf}}{d\tau_h}+\frac{dp_{hh}}{d\tau_h}q_{hh}+(p_{hh}-c)\frac{dq_{hh}}{d\tau_h}. \end{aligned} \tag{37}$$

Notice that markets are segmented such that production decisions in Foreign are independent of those in Home, or $\frac{dp_{fh}}{d\tau_h}=0$, $\frac{dq_{fh}}{d\tau_h}=0$ and $\frac{d\tau_f}{d\tau_h}=0$. I next substitute the demand functions in (2) into (37) which, after some algebraic manipulations, reduces (37) to:

$$\frac{dW_h}{d\tau_h}=-q_{hf}\frac{dp_{hf}}{d\tau_h}+q_{hf}+\tau_h\frac{dq_{hf}}{d\tau_h}+(p_{hh}-c)\frac{dq_{hh}}{d\tau_h}.$$

Defining the net-of-tariff price of the Foreign good sold in the Home market as $p^*_{hf}=p_{hf}-\tau_h$ I obtain the expression in (10). 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 derive the analogous decomposition of the welfare effects in that country. **q.e.d.**

To find the effect of the Foreign tariff on Home welfare I take the derivative of (36) wrt. $\tau_f$. This yields:

$$\frac{dW_h}{d\tau_f}=\frac{dp_{fh}}{d\tau_f}q_{fh}-q_{fh}+(p_{fh}-c-\tau_f-\alpha_f)\frac{dq_{fh}}{d\tau_f}. \tag{38}$$

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 (14). **q.e.d.**

**Derivation of Eq. (11), Eq. (13) and Eq. (15)**

Substituting the inverse demand functions in (2) and the Cournot quantities in (5) into (10), yields the following expression:

$$\begin{aligned} \frac{dW_h}{d\tau_h} &= \frac{(\Gamma(0,\gamma))^2-\Gamma(0,\gamma)\,\Gamma(2,\gamma)\,(\tau_h+\alpha_h)}{(\Gamma(0,\gamma)\,\Gamma(2,\gamma))^2} \\ &\quad-\tau_h\frac{2}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)}+\frac{\gamma}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)} \\ &\quad-\left[\gamma\frac{\Gamma(0,\gamma)-2(\tau_h+\alpha_h)}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)}\right]\frac{\gamma}{\Gamma(0,\gamma)\,\Gamma(2,\gamma)}. \end{aligned} \tag{39}$$

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 (11). 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 (39) wrt. $\alpha_h$, I obtain the expression in (13). **q.e.d.**

By writing up expressions for the inverse demand functions and Cournot quantities in the Foreign market, an expression for (38) can be found as:

$$\frac{dW_h}{d\tau_f} = -\frac{\Gamma(0,\gamma) - 2(\tau_f + \alpha_f)}{\Gamma(0,\gamma)\Gamma(2,\gamma)}\frac{4}{\Gamma(0,\gamma)\Gamma(2,\gamma)}. \tag{40}$$

Substituting the Foreign Nash tariff from (11) into (40) yields the expression in (15). **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 (11). 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 (11) 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. (17), Eq. (18) and Eq. (20)**

The joint welfare of Home and Foreign is given as:

$$\frac{1}{1-\delta}J(\tau_h,\tau_f,\alpha_h,\alpha_f) = \frac{1}{1-\delta}\left(W_h(\tau_h,\tau_f,\alpha_h,\alpha_f) + W_f(\tau_h,\tau_f,\alpha_h,\alpha_f)\right). \tag{41}$$

The expression for welfare in Home is obtained by substituting (7) and (8) into the expression for Home welfare (9). The equivalent expression for Foreign welfare is easily obtained by writing up the equivalent welfare expression by exchanging $h$ and $f$ in (9). Substituting the expressions for Home and Foreign welfare into (41) yields:

$$\begin{aligned} J\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right) &= CS_h\left(\tau_h,\alpha_h\right)+TR_h\left(\tau_h,\alpha_h\right)+\Pi_h\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right) \\ &\quad +CS_f\left(\tau_f,\alpha_f\right)+TR_h\left(\tau_f,\alpha_f\right)+\Pi_f\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right) \\ &= \frac{1}{2}\left(a-p_{hh}\right)q_{hh}+\frac{1}{2}\left(a-p_{hf}\right)q_{hf}+\tau_h q_{hf} \\ &\quad +\left(p_{hh}-c\right)q_{hh}+\left(p_{fh}-c-\tau_f-\alpha_f\right)q_{fh} \\ &\quad +\frac{1}{2}\left(a-p_{ff}\right)q_{ff}+\frac{1}{2}\left(a-p_{fh}\right)q_{fh}+\tau_f q_{fh} \\ &\quad +\left(p_{ff}-c\right)q_{ff}+\left(p_{hf}-c-\tau_h-\alpha_h\right)q_{hf}. \end{aligned} \tag{42}$$

Differentiating (42) with respect to $\tau_h$ yields

$$\begin{aligned} \frac{dJ}{d\tau_h} &= -\frac{1}{2}\frac{dp_{hh}}{d\tau_h}q_{hh}+\frac{1}{2}\left(a-p_{hh}\right)\frac{dq_{hh}}{d\tau_h}-\frac{1}{2}\frac{dp_{hf}}{d\tau_h}q_{hf}+\frac{1}{2}\left(a-p_{hf}\right)\frac{dq_{hf}}{d\tau_h} \\ &\quad +q_{hf}+\tau_h\frac{dq_{hf}}{d\tau_h}+\frac{dp_{hh}}{d\tau_h}q_{hh}+\left(p_{hh}-c\right)\frac{dq_{hh}}{d\tau_h}+q_{hf}\frac{dp_{hf}}{d\tau_h} \\ &\quad -\left(\tau_h+\alpha_h\right)\frac{dq_{hf}}{d\tau_h}+\left(p_{hf}-c\right)\frac{dq_{hf}}{d\tau_h}. \end{aligned} \tag{43}$$

Substituting the inverse demand functions (2) and the equivalent demand functions for Foreign into (43) and performing several algebraic steps reduces (43) to:

$$\frac{dJ}{d\tau_h}=\left(p_{hh}-c\right)\frac{dq_{hh}}{d\tau_h}+\left(p_{hf}-c\right)\frac{dq_{hf}}{d\tau_h}-\alpha_h\frac{dq_{hf}}{d\tau_h}, \tag{44}$$

which is the expression in (17). Substituting the inverse demand functions (2) and the Cournot quantities (5) and the equivalent functions for Foreign into (44) yields:

$$\begin{aligned} \frac{J\left(\tau_h,\tau_f,\alpha_h,\alpha_f\right)}{d\tau_h} &= \frac{-2\Gamma\left(0,\gamma\right)\Gamma\left(0,\gamma\right)+\alpha_h\left(4+\gamma^2\right)+\alpha_f\left(4+\gamma^2\right)-\tau_h\left(8-6\gamma^2\right)}{\left(\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right)^2} \\ &= \frac{-2\Gamma\left(0,\gamma\right)\Gamma\left(0,\gamma\right)+\alpha_h\left(4+\gamma^2\right)+\alpha_f\left(4+\gamma^2\right)}{\left(\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right)^2}-\frac{\tau_h\left(8-6\gamma^2\right)}{\left(\Gamma\left(0,\gamma\right)\Gamma\left(2,\gamma\right)\right)^2}. \end{aligned} \tag{45}$$

Solving for $\tau_h$ yields:

$$\tau_h^E=\frac{-\Gamma\left(0,\gamma\right)\Gamma\left(0,\gamma\right)+\frac{1}{2}\alpha_h\left(4+\gamma^2\right)+\frac{1}{2}\alpha_f\left(4+\gamma^2\right)}{\left(4-3\gamma^2\right)}. \tag{46}$$

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 (46) equal to (11) I obtain:

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

Solving for $\alpha$ yields $\alpha = 1 - \frac{3}{4}\gamma$, which is the upper bound proposed in (19). **q.e.d.**

**Derivation of Eq. (25), Eq. (29) and Eq. (30)**

In order to derive the level of patience required for an agreement based on the efficient tariffs in (18) 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 (9), and then evaluate them at the various tariff levels.Thus, after substantial algebraic manipulations, I obtain:

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

Substituting these expressions into (24) yields (25). **q.e.d.**

The minimum self-enforceable tariffs $\left(\tau_h^M,\tau_f^M\right)$ can be found by solving (28). 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 (28) using the lagrange method:

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

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:

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

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

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

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

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

$$\begin{aligned}\lambda_1 &= \lambda_2 = \frac{16\delta + 3\gamma^2 - 6\gamma^2\delta - 12}{\delta(1+\delta)(12 - 8\delta - 3\gamma^2)};\\ \tau_h &= \tau_f = \frac{\alpha(40\delta + 3\gamma^2 - 12) - 40\delta - 3\gamma^2 + 12 + 24\delta\gamma}{3(12 - 3\gamma^2 - 8\delta)}.\end{aligned} \tag{47}$$

It is easy to show that the minimum enforceable tariffs are equal to the Nash tariffs in the limit. Setting (47) equal to (11) yields:

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

Solving for $\alpha$ yields $\alpha = 1 - \frac{3}{4}\gamma$, which is the uppoer bound proposed in (12). **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 (7) and (8) into (9) yields an expression for Home welfare. Evaluating this when Home plays Nash by imposing the Nash tariff in (11), 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:

$$W_h^D\left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right) - W_h^C\left(\tau_h^C, \tau_f^C, \alpha_h, \alpha_f\right) = \frac{(1-\alpha_h)^2}{6(2-\gamma)(2+\gamma)}. \tag{48}$$

Differentiating wrt. $\alpha_h$ yields:

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

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:

$$W_h^D\left(\tau_h^N, \tau_f^C, \alpha_h, \alpha_f\right) - W_h^P\left(\tau_h^N, \tau_f^N, \alpha_h, \alpha_f\right) = \frac{4(1-\alpha_f)(5(1-\alpha_f) - 3\gamma)}{9(2-\gamma)^2(2+\gamma)^2}. \tag{49}$$

Differentiating wrt. $\alpha_f$ yields:

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

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 (48) by (49) yields an expression for the critical discount factor for symmetric trade costs:

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

which is the expression in (33). Taking the derivative wrt. $\alpha$ yields:

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

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


[1] Baier, S. L., and Bergstrand, J. H. 2004. Economic Determinants of Trade Liberalisation. *Journal of International Economics*. 64(1), 29-63.

[2] Baldwin, R.E., and Venables, A.J. 1995. Region Economic Integration, in *Handbook of International Economics,* Elsevier. Ed. by Grossman, G.M., and Rogoff, K., vol. 3, Chap. 31, pp. 1597-1644.

[3] Bond, E., Syropoulos, C. 1996. Trading Blocs and the Sustainability of Inter-Regional Cooperation, in *The New Transatlantic Economy*. London: Cambridge University Press. Ed. by Conzoneri, M., Ethier, W., Grilli, V., 118-40.

[4] Bond, E. 2001. Multilateralism vs Regionalism: Tariff Cooperation and Interregional Transport Costs, in *Regionalism and Globalisation: Theory and Practice*. London Routledge. Ed. by Lahiri, S.,16-38.

[5] Bond, E., Syropoulos, C., Winters, A. 2001. Deepening of Regional Integration and Multilateral Trade Agreements, *Journal of International Economics,* 53(2), 335-61.

[6] Brander, J. A. 1981. Intra-industry Trade in Identical Commodities. *Journal of International Economics*, 11(1), 1-14.

[7] Furusawa, T. and Lai, E. L-C. 1999. Adjustment Costs and Gradual Trade Liberalization. *Journal of International Economics*, 49(2), 333-361.

[8] Grossman, G. M., and Helpman, E. 1995. Trade Wars and Trade Talks, *Journal of Political Economy*, 103(4), 675-708.

[9] Johnson, H. G., 1954. Optimum Tariffs and Retaliation, *Review of Economic Studies*, 21(2), 142-53.

[10] Ludema, R. D. 2002. Increasing Returns, Multinationals and Geography of Preferential Agreements. *Journal of International Economics*, 56(2), 329-358.

[11] Maggi, G. and Rodriguez-Clare, A. 2007. A Political-Economy Theory of Trade Agreements. *American Economic Review*, 97(4), 1374-1406.

[12] McGowan, D., and Milner, C. 2011. Trade Costs and Trade Composition. Unpublished.

[13] Mrazova, M. 2011. Trade Agreements when Profits Matter. Unpublished.

[14] Novy, D. 2010. Gravity Redux: Measuring International Trade Costs with Panel Data, *Warwick Economic Research Papers*.

[15] Venables, A. J., Trade and Trade Policy with imperfect competition: the case of identical products and free entry, *Journal of International Economics*, 19(1-2), 1-19.

[16] Venables, A. J., Trade and Trade Policy with Differentiated Products: A Chamberlinian-Ricardian Model, *Economic Journal*, 97(387), 700-717.

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

[18] Zissimos, B. 2011. Why are Trade Agreements Regional?, *Review of International Economics*, 19(1), 32-45.