Revista de Economía
del Rosario

The endogenous formation of sustainable trade agreements

Mette Ersbak Bang Nielsen*
Department of Economics, Princeton University, USA.

Received: June 2005 – Approved: July 2005

Abstract. This paper addresses the endogenous formation of trade agreements in a three-country model of imperfect competition. While the requirement of sustainability of preferential trade areas has often been ignored in the literature, I construct a framework for predicting which trade agreements form when sustainability is explicitly included as a constraint on the formation of cooperative agreements. It is found that the introduction of a self-enforcement requirement reduces the overall scope for cooperative trade agreements, and that preferential trade areas can be stepping stones or stumbling blocks depending on the size of relative demand between countries.

Key words: trade agreements, customs unions, FTAs.
JEL classification: F13, F15.

Resumen. Este artículo se refiere a la formación de acuerdos endógenos de libre comercio en un modelo de competencia imperfecta entre tres países. Teniendo en cuenta que la literatura ha ignorado en ocasiones el requisito de sostenibilidad en las áreas de libre comercio, el artículo propone una estructura para predecir qué acuerdos comerciales se forman cuando la sostenibilidad es explícitamente incluida como una restricción en la formación de acuerdos cooperativos. Se encuentra que la introducción de un requisito de auto-ejecución reduce el alcance general de los acuerdos comerciales y que las áreas comerciales preferenciales pueden ser *stepping stones* o *stumbling blocks*, dependiendo de la demanda relativa entre los países.

Palabras clave: acuerdos comerciales, uniones aduaneras, ALCs.
Clasificación JEL: F13, F15.

*This paper is based on the 2004 University of Warwick MSc in Economics dissertation of the author. I am grateful to Carlo Perroni for helpful comments and guidance.

Address for correspondence: Fisher Hall 100, Department of Economics, Princeton University, Princeton, NJ-08544, USA.
E-mail: mnielsen@princeton.edu

1. Introduction

Between 1948 and 1994, the GATT received 124 notifications of regional trade agreements (RTAs) related to trade in goods. In the first seven years after the creation of the WTO in 1995, the notification of more than 130 additional agreements covering trade in goods and services followed. In 2003, more than 150 RTAs were in force, and the number was expected to rise to around 300 in 2007 if the trend then continued. The vast majority of these agreements are either customs unions (CUs) or free trade areas (FTAs).[1]

The proliferation of preferential trade agreements (PTAs) implies the presence of major incentives to enter such agreements. A growing literature tries to identify these incentives, explain how the formation of PTAs affect welfare of members as well as non-members, and predict the impact of PTAs on the organization of the world trading system. This literature can broadly be divided into two approaches.[2] The traditional one follows Viner (1950) and examines the effect of exogenously specified PTAs on the welfare of members and non-members.[3] The second, and more recent, approach models the endogenous formation and stability of PTAs in order to understand how the pattern of trade agreements can be expected to develop over time. Especially interesting is the question of whether PTA formation will speed up the process towards global free trade or end up blocking the process in order to preserve preferential gains. The stepping stone/ stumbling block idea was introduced by Bhagwati (1991), and has been the topic of several papers including Ludema (1994), Bagwell and Staiger (1997a, b), Yi (1996), and Krishna (1998).

This paper emphasizes the latter approach in that it provides a first attempt to predict which trade agreements result when the formation of PTAs (and free trade) is endogenous and all agreements must be self-sustaining. In particular, I determine the welfare effects of eight different trade agreements (namely the non-cooperative, 3 FTAs, 3 CUs, and free trade)[4] in a three-country model of imperfect competition and endogenous tariff setting. I use welfare comparisons to determine the individual countries' preferences over trade agreements and find conditions for agreements to be sustainable in a standard infinitely repeated game framework with Nash punishment. The core solution concept developed by Riezman (1985) is applied to the model, and the idea expanded to a dynamic framework in which trade agreements are restricted to those that are sustainable and therefore will be observed: while Riezman assumes that "some" mechanism will make the trade agreements of the core binding, I explicitly

[1] www.wto.org/english/tratop_e/region_e/regfac_e.htm#top
and www.warwick.ac.uk/~vxscbl/regionalism.pdf

[2] Yi (1996) p. 154.

[3] Contributions include Kemp and Wan (1976), Panagariya and Krishna (2002), and Krugman (1991).

[4] I limit my analysis to these agreements since the vast majority of PTAs in the world are either FTAs or CUs. Though the inclusion of multilateral tariff reductions would be desirable, it is beyond the scope of this paper except for a brief discussion in the case of symmetric countries. With this limitation, PTAs are subsequently taken to refer to CUs and FTAs only.

model sustainability in order to make everything endogenous, self-sustainable and optimal. These requirements reflect the facts that countries voluntarily choose to make a trade agreement, that they cannot prevent other countries from cooperating, and that real world trade agreements face enforcement problems which should be dealt with theoretically. An important methodological contribution of the paper is to introduce the core as an equilibrium selection tool which helps determine the trade agreement(s) most likely to form among all self-sustainable agreements.

One example of how self-enforcement cannot be taken for granted is the Andean Community. It started in 1969 with the signature of the Cartagena Agreement between Bolivia, Chile, Colombia, Ecuador, and Peru, aiming at the creation of a CU. Though successful at first, problems started in the late 1970's when members failed to comply with the agreement due to pressure from domestic interest groups "unhappy with the differences in short-term costs and benefits flowing from the integration process".[5] Unresolved disputes along with the oil crisis worsened the situation, and the Pact was virtually dead until a fundamental change towards free market-oriented policies was adopted in the early 1990's. Since then, integration has enhanced. First with an FTA structure but then with gradual integration towards a CU.[6]

Naturally, the failure of the Community in the 1980's was not only due to Latin countries being "too impatient". However, the attempts to start and later restore the Community do imply an incentive to cooperate which history tells us did not lead to cooperation. In this way, the example shows that the sustainability of PTAs should not simply be assumed, and that an incentive to cooperate does not always equal the ability to do so.

As for multilateral agreements, enforcement has improved significantly in later years through the dispute settlement body of the WTO. But since a dispute can still take more than one year to settle, countries continue to have an incentive to cheat to obtain short term gains- and still do cheat as has been the case with the US recently.

The idea of using core theory to predict which trade agreement form when countries are free to make agreements endogenously was developed in Riezman (1985). He considers a three country- three goods pure exchange economy where countries differ in endowments only. CUs are seen as coalitions in the core, and the formation of these modelled in a two stage game. First, a coalition is chosen based on core theory. Second, optimal tariffs are chosen given the coalition. The resulting agreements are stable in the sense that no other coalition which yields higher welfare to all members of that coalition can form. It is assumed that there exists a mechanism so that agreements are binding. This gives a natural interpretation of the core, namely that agreements in the core will be observed.

[5] www.mercosurconsulting.net/Articles/article2.html

[6] www.ictsd.org/dlogue/1999-02-01/overviewAndean.pdf and www.mercosurconsulting.net/Articles/article2.html

Riezman's idea was subsequently expanded in Krishna (1998), which considers a three-country oligopolistic trade model with several firms in each country, and takes a political economy approach where producer lobbies' preferences are determinant for trade policy. It is found that the more trade diverting an FTA is, the more likely it is to be supported by its members because trade diversion towards products produced by members move the game from being close to zero-sum (members gain in their partner's market but lose in their own) to being positive sum (members gain more in their partner's market and lose less at home at the expense of the third country). It is also found that the more trade diverting a block is, the more likely it is to be a stumbling block to free trade.

While the above coalition formation games are static, Freund (2000) expands the model of Krishna (1998) to a dynamic framework to show that multilateral tariff reductions (which are assumed binding) increase both the incentive for PTAs and the likelihood that these are sustainable. Sustainability for PTAs is, as here, defined in a standard repeated game framework. The incentive for PTAs is shown to depend on the relative size of two effects: the efficiency effect of free trade- the gains from the increase in competition following full trade liberalization-, and the redistributive effect of PTAs- the diversion of profit loss to the outsider country, whose output contracts when the PTA is formed.

Finally, Yi (1996) studies the stability of CU structures in an N country model under two different rules: "open regionalism" under which coalitions can form freely provided no outsider is excluded, and "unanimous regionalism" under which an outsider country can join a CU iff all existing members agree to this. He finds that open regionalism supports free trade, while unanimous regionalism typically supports two CUs of asymmetric size in equilibrium. Furthermore, CUs are stepping stones towards free trade under open regionalism, but can be stumbling blocks under unanimous regionalism.

These papers all yield interesting results, but what is also evident is that at least some coalitions, once made, are assumed binding. As noted by Syropoulos (1999), this approach is customary in the literature. Indeed, Baldwin, in his infamous (1995) domino-paper does not account for why the open regionalism rule is applicable. Bagwell and Staiger (1997a, b) argue that enforcement is important and require multilateral agreements to be sustainable but still let CUs form exogenously and expand at random. Krugman (1991) considers symmetric CUs, but does not account for their sustainability. The present paper shows that this approach can be misleading in that predictions of which trade agreements form change significantly by the imposition of the sustainability requirement.

The paper is organized as follows: the static model is presented in section 2 along with the calculation of optimal tariffs under all possible trade agreements. Section 3 introduces the static game and the welfare incentives that drive it, and determine the trade agreements in the core. Section 4 defines sustainability and solves the dynamic game. The stepping stone/ stumbling block issue is briefly touched upon, and the model is extended to cover multilateral tariffs

when countries are symmetric. Section 5 concludes.

2. The static model

This section outlines the underlying model with optimizing consumers and producers. The welfare measure is presented, and by optimizing it, optimal tariffs are determined for the uncooperative equilibrium as well as for preferential agreements.

The basic model builds on Krishna (1998) and Freund (2000). It contains three countries: X, Y, and Z, each with one profit maximizing firm producing a homogenous, imperfectly competitive good. Segmented markets characterized by Cournot competition lead to trade in this good, with the three profit maximizing firms competing in all markets (countries). The good is traded without transportation costs,[7] but possibly subject to an above-zero specific import tariff. In consensus with recent negotiations within the WTO, export subsidies are banned.

The economy furthermore contains a numeraire good which is characterized by perfect competition and serves to settle the trade balance. No tariff is levied on this good and transporting it between countries is costless.

Consumers in each country maximize their utility, which is quasilinear in the two goods:

$$U_x = \xi_x + a_x q_x - \frac{1}{2} b_x q_x^2 \tag{1}$$

with ξ_x and q_x denoting consumption in country X of the numeraire and imperfectly competitive good respectively. Here, and throughout this and the next section, country (firm in, consumers in, etc.) X will be used as an example for notational ease. Results for Y and Z follow directly by substitution of appropriate sub- and superscripts due to the symmetric nature of the model. a and b are parameters > 0, allowed to be country specific (therefore the subscript). Maximizing utility gives a linear inverse demand function for the imperfectly competitive good:

$$p_x(q_x) = a_x - b_x q_x \tag{2}$$

Firms maximize profits taking demand, tariffs, the supply of other firms, and production costs as given. The latter are constant on the margin (in terms of the numeraire good), so total costs for the firm from country X (firm X henceforth) are given by $C_x = c_x q^x$ where c_x is the marginal cost of firm X ($c_x < a_x$), and q^x is the total production of that firm. Along with the assumption of segmented markets, this allows me to treat the profit maximization in each market separately.[8]

[7] As in Krishna (1998), the joint assumption of segmented markets and no transportation costs on direct export of the imperfectly competitive good can be justified by assuming that transportation costs are prohibitively high for any third party trying to make an arbitrage profit.

[8] For example, this would not hold had their been economies of scale, because the marginal

The profit maximization problem of firm X in country i ($i \in (X, Y, Z)$ throughout this paper) is then given by:

$$\begin{aligned} &\underset{x_i}{\text{Max}} \{\pi_i = x_i p_i(q_i) - c_x x_i - t_i^x x_i\} \\ &\text{s.t.} \quad y_i;\ z_i;\ p_i(q_i) = a_i - b_i q_i;\ q_i = x_i + y_i + z_i \end{aligned} \tag{3}$$

where π_i is profit of firm X in country i; x_i, y_i, z_i is the supply of firm X, Y, Z respectively in country i, and t_i^x is the tariff levied on imports from firm X by country i. Note how tariffs simply increase the constant marginal cost of getting each unit to the market: the *effective* marginal cost becomes $c_x + t_i^x$ instead of c_x.

To solve the maximization problem, assume that costs and demands are related in a way that yields interior solutions only (that is, above zero equilibrium quantities and profits). Then solving the maximization problem yields a supply of firm X in country i of

$$x_i^* = \frac{1}{2b_i}(a_i - c_x - t_i^x) - \frac{1}{2}(y_i + z_i) \tag{4}$$

As expected, supply is increasing in demand $\left(\text{i.e. in } a_i \text{ and } \frac{1}{b_i}\right)$ and decreasing in the cost of production and tariffs levied against the product. It is also decreasing in the supply of other firms in market i since this lowers the residual demand curve facing firm X.

In equilibrium, supply equals demand and there are neither trade deficits nor surpluses. In this model, equilibrium quantities can be found directly from solving the nine supply equations (i.e. the 9 equations resulting from (4) when one considers supply from each country to each country) since prices will adjust to create demand-supply equalization and the trade balance is settled automatically due to the presence of the numeraire good: utility is increasing in ξ, so countries will buy exactly as much ξ from abroad as their trade balance allows them to. Equilibrium consumption in country X is then:

$$q_x = \frac{1}{4b_x}(3a_x - c_x - c_y - c_z - t_x^y - t_x^z) \tag{5}$$

which is increasing in country X's demand and decreasing in effective production costs of all firms because supply is (cf. (4)). The reason why consumption in X is independent of demand in Y and Z is that each firm's supply to country X is independent of its supply to other markets. On the other hand, all supply decisions involve considerations of marginal cost, and consumption therefore

cost of export to another country then depends on the exporting firm's total production and thus on the conditions in other markets. Furthermore, the decrease in production costs following an expansion of output would create an additional incentive for cooperation among countries which is ignored here.

More generally, any factor that causes interdependency between markets could cause a link between production/ sales decisions in different markets. Dixit (1987) provides some examples of this.

depends on the (effective) cost of production in each firm. The total production for market X is divided among firms in X, Y, Z as follows:

$$\begin{aligned} x_x &= \frac{1}{4b_x}\left(a_x - 3c_x + c_y + c_z + t_x^y + t_x^z\right) \\ y_x &= \frac{1}{4b_x}\left(a_x + c_x - 3c_y + c_z - 3t_x^y + t_x^z\right) \\ z_x &= \frac{1}{4b_x}\left(a_x + c_x + c_y - 3c_z + t_x^y - 3t_x^z\right) \end{aligned} \tag{6}$$

These are simply the reduced forms of the 3 supply equations similar to (4). By differentiation,

$$\frac{dx_x}{dt_x^y} = \frac{dx_x}{dt_x^z} = \frac{1}{4b_x}, \quad \frac{dy_x}{dt_x^y} = \frac{dz_x}{dt_x^z} = -\frac{3}{4b_x} \quad \text{and} \quad \frac{dy_x}{dt_x^z} = \frac{dz_x}{dt_x^y} = \frac{1}{4b_x} :$$

the production for firms' own markets is increasing in the degree of protection against imports, while exports are decreasing in the tariffs levied against the exporting country but increasing in tariffs levied against other import.[9] The intuition of these results should be clear: the more protected a firm is in a given market relative to other firms, the more the firm will produce for that market.

2.1. Welfare

The welfare criterion used by the government is an equally weighted function of consumers surplus, producers surplus, and tariff revenue. The equal weighting may reflect either the reduced form of a particular political economy game, or an abstraction from political economy altogether.[10] Tariff revenue is included in the welfare function as revenue gives the country access to a larger import of the numeraire good ξ. For country X, I thus obtain:

$$\begin{aligned} W_x\left(t_x^y, t_x^z, t_y^x, t_y^z, t_z^x, t_z^y\right) &= CS_x + PS_x + TR_x \\ &= \frac{1}{2}\big(a_x - p_x(q_x)\big)q_x + \sum_{i=x,y,z} x_i\big(p_i(q_i) - c_x - t_i^x\big) + t_x^y y_x + t_x^z z_x \end{aligned} \tag{7}$$

[9] Krishna (1998) reaches a similar result for his model with more than one firm in each country.

[10] Macho-Stadler et al (1998) show how different weights in the welfare function can affect the stability of trade agreements in a similar model.

By substituting the equilibrium quantities obtained previously, this reduces to:

$$\begin{aligned} W_x\left(t_x^y, t_x^z, t_y^x, t_y^z, t_z^x, t_z^y\right) = & \frac{1}{32b_x}\left(3a_x - c_x - c_y - c_z - t_x^y - t_x^z\right)^2 \\ & + \frac{1}{16b_x}\left(a_x - 3c_x + c_y + c_z + t_x^y + t_x^z\right)^2 \\ & + \frac{1}{16b_y}\left(a_y - 3c_x + c_y + c_z - 3t_y^x + t_y^z\right)^2 \\ & + \frac{1}{16b_z}\left(a_z - 3c_x + c_y + c_z - 3t_z^x + t_z^y\right)^2 \\ & + \frac{1}{4b_x}\left(a_x + c_x - 3c_y + c_z - 3t_x^y + t_x^z\right)t_x^y \\ & + \frac{1}{4b_x}\left(a_x + c_x + c_y - 3c_z - 3t_x^z + t_x^y\right)t_x^z \end{aligned} \tag{8}$$

with the six lines each corresponding respectively to consumers surplus, producers surplus in X, Y and Z, and tariff revenues obtained from X's imports from Y and Z.

It is seen that consumers surplus decreases when demand falls and when production costs or protection rise. This is intuitive since these changes all reduce the quantity supplied, and this lowers consumer surplus as evident in Figure 1 below. All profits are increasing in demand and the (effective) cost advantage of firm X. Finally, the tariff revenue from Y has a "Laffer curve" effect: an increase in tariffs decreases the tax base but increases the per unit revenue so that revenue is increasing in t_x^y for $t_x^y < \frac{1}{6}\left(a_x + c_x - 3c_y + c_z + t_x^z\right)$, and decreasing in t_x^y otherwise. Tariff revenues from country Z change in a similar manner.

To get an overview of how the surplus of country X is divided between countries, consider Figure 1. In the figure, the downwards sloping curve is the demand in X, the horizontal axis measures q_x, and the vertical axis measures p_x. The total gain for country X in its own market is given by consumers surplus (AEB) + producers surplus (DELK) + tariff revenue from Y (FGNM) + tariff revenue from Z (HIJO). This is the area that country X maximizes when setting its external tariffs below. The remaining surplus goes to firms Y (area BCGF) and Z (area CDIH) as profits.

2.2. Optimal tariffs

Tariffs are set by the government of each country to maximize welfare. As in Riezman (1985), the choice of trade agreement and the setting of optimal tariffs take place in a two stage game. Optimal tariffs are set in the second stage, conditioned on the relevant trade agreement (free trade, FTA, CU, or Nash). This trade agreement is determined in the first stage, with each country

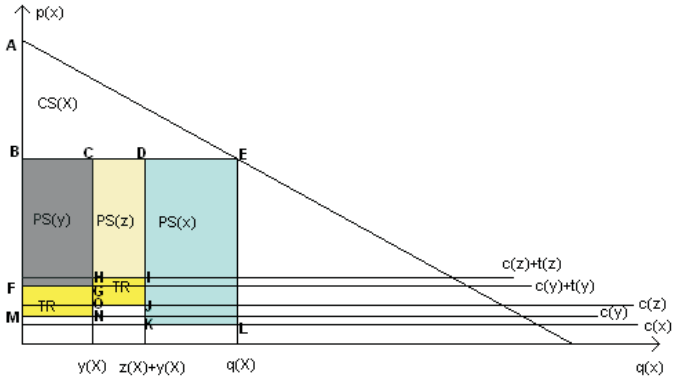


Figure 1. Division of gains from country X.

seeking to maximize its welfare, taking into account the choice of optimal tariffs in the second stage.

Both the parameters that the government can determine and the welfare outcome depends on the actual trade agreement. For the home country, increasing an external tariff will cause a decrease in consumers surplus (q_x falls), an increase in producers surplus at home (competition from abroad decreases), and have an ambiguous effect on the tariff revenue (there is less import, but each unit of import generates a larger revenue). Producers surplus abroad goes unaffected because of market segmentation.

Governments will set the tariff at the point where the gain from increasing it any further is exactly matched by the loss of doing so. That is, where the derivatives of the welfare function with respect to relevant tariffs are zero. While country X is still used as an example, optimal tariffs for all countries can easily be determined by exchanging appropriate sub- and superscripts.

2.2.1. Nash equilibrium tariffs

The Nash equilibrium tariffs are the tariffs of the non-cooperative equilibrium. Because countries act individually, they are free to optimize welfare (8) with respect to two external tariffs. For country X, this optimization yields the first order conditions:

$$\frac{dW_x\left(t_x^y, t_x^z, t_y^x, t_y^z, t_z^x, t_z^y\right)}{dt_x^y} = \frac{dW_x\left(t_x^y, t_x^z, t_y^x, t_y^z, t_z^x, t_z^y\right)}{dt_x^z} = 0$$

which yields

$$\begin{aligned} t_x^y &= \frac{1}{20}\left(6a_x - 2c_x - 7c_y + 3c_z\right) \\ t_x^z &= \frac{1}{20}\left(6a_x - 2c_x + 3c_y - 7c_z\right) \end{aligned} \tag{9}$$

In the Nash (and also under subsequent trade agreements), optimal tariffs of any one country are independent of the tariffs set by other countries. This lack of strategic interaction results from the assumed lack of interaction between markets. Had preferences for instance not been quasilinear and income effects of tariffs on the imperfectly competitive good therefore not effectively assumed away, these effects would cause interdependency in tariff setting across countries. The lack of strategic interaction is desirable in that it makes the model tractable. However, it must also be recognized that potentially important effects are left out of the analysis, and that results should be interpreted with care. Of course, as is evident from the welfare function (8), the lack of strategic interaction in tariff setting by no means imply that welfare is also independent of other countries' tariffs.

The optimal tariffs are increasing in a_x: when a_x increases, country X is willing to pay a higher price for each quantity (remember $p_x(q_x) = a_x - b_x q_x$), and therefore foreign supply rises and more profits go abroad $\left(\text{it can be shown that the profit of firm } X \text{ in country } i \text{ equals } PS_i = b_i(x_i)^2\right)$. The government in country X can shift some of these profits back by raising tariffs. As it turns out, a small increase in tariffs benefits the home firm and tariff revenue more than the resulting drop in foreign supply harms the consumers in X. Thus, it is optimal to raise external tariffs in face of an increase in domestic demand. Costs affect the optimal tariffs for similar reasons. For instance, the larger c_y the smaller is Y's market share in X, and the bigger is the competition from Z's exports, which calls for a larger tariff towards Z.

2.2.2. FTA tariffs

In conformity with the requirements of the GATT Article XXIV, all PTAs are restricted to have internal free trade. FTAs and CUs are differentiated by their external tariff setting: FTAs set their external tariffs independently while CUs set a common external tariff (CET) which maximizes the joint welfare of members. It is assumed that each member in the PTA obtains all tariff revenue associated with its consumption. Thus, in an XY-PTA,[11] X acquires the tariff revenue obtained from the import of z_x, and Y obtains the tariff revenue obtained from importing z_y. No intercountry transfers are allowed for.[12]

Let X form an FTA with Y. Under the above conditions, t_x^y is fixed at zero and X can only maximize its welfare with respect to its tariff against the outsider country (Z). Calculating

$$\frac{d}{dt_x^z} W_x\left(0, t_x^z, 0, t_y^z, t_z^x, t_z^y\right) \quad \text{and} \quad \frac{d}{dt_y^z} W_y\left(0, t_x^z, 0, t_y^z, t_z^x, t_z^y\right)$$

[11] That is, a PTA between X and Y. Similar notation will be used throughout the paper.

[12] Kowalzcyk and Sjostrom (1994) show how transfers can be used to facilitate trade negotiations. Riezman (1985) briefly explains one complication that can arise when transfers are allowed for.

and equating the derivatives to zero yields Proposition 1:

Proposition 1. *The unique optimal external tariffs of FTA-members X and Y are*

$$t_x^z = \frac{1}{21}(3a_x - c_x + 7c_y - 9c_z) \tag{10}$$

$$t_y^z = \frac{1}{21}(3a_y + 7c_x - c_y - 9c_z) \tag{11}$$

The external tariffs of the FTA are (normally) below those of the Nash[13] due to what has been named the "tariff complementarity" effect: when a member country removes the tariff against the partner, imports from the partner expand while imports from the non-member contract (due to the lack of preferential status). The latter lowers both consumers surplus and tariff revenue. As noted by Freund (2000) p. 364, these negative effects on welfare may be reduced by slightly lowering the tariff against the non-member.

The optimal tariff of X is still increasing in a_x and c_y while decreasing in c_x and c_z for reasons similar to those under the Nash. The non-member country is free to optimize welfare with respect to its two external tariffs and will therefore set the same tariffs as under the Nash.

2.2.3. CU tariffs

In a CU, internal trade is free and the CET set by maximization of the joint welfare of member countries: letting $t_{x,y}^z$ denote the CET of the CU between X and Y, $t_{x,y}^z$ must satisfy

$$\frac{d}{dt_{x,y}^z}\left(W_x\left(0, t_{x,y}^z, 0, t_{x,y}^z, t_z^x, t_z^y\right) + W_y\left(0, t_{x,y}^z, 0, t_{x,y}^z, t_z^x, t_z^y\right)\right) = 0.$$

Because the effect on the partner's profits is internalized, the optimal CET is usually greater than the external tariffs of the corresponding FTA:[14]

Proposition 2. *The unique optimal CET of CU-members X and Y is:*

$$t_{x,y}^z = \frac{5\left(b_y a_x + b_x a_y\right) + \left(b_y + b_x\right)\left(c_x + c_y - 7c_z\right)}{19\left(b_y + b_x\right)} \tag{12}$$

[13] For country X, the condition for this is $66a_x - (22c_x + 77c_y - 33c_z) > 0$.

[14] This is true when

$$b_x\left(-57a_x + 105a_y + 40c_x - 112c_y + 24c_z\right) + b_y\left(48a_x + 24c_z + 40c_x - 112c_y\right) > 0.$$

By differentiation,

$$\frac{dt^z_{x,y}}{da_x} > 0; \quad \frac{dt^z_{x,y}}{da_y} > 0;$$

$$\frac{dt^z_{x,y}}{dc_x} > 0; \quad \frac{dt^z_{x,y}}{dc_y} > 0; \quad \frac{dt^z_{x,y}}{dc_z} < 0;$$

$$\frac{dt^z_{x,y}}{db_x} = \frac{5}{19}\frac{(a_y - a_x)\, b_y}{(b_y + b_x)^2}; \quad \frac{dt^z_{x,y}}{db_y} = \frac{5}{19}\frac{(a_x - a_y)\, b_x}{(b_y + b_x)^2}.$$

The main difference to the FTA is that the optimal tariff is now increasing in demand and production cost of both X and Y as a result of the internalization of effects on the partner's welfare. Also, the optimal tariff is increasing in b_x and decreasing in b_y provided $a_y > a_x$, and vice versa when $a_y < a_x$: if the intersection of demand in country Y is greater than in X, then a decrease in demand in X (that is, an increase in b_x) calls for more protection.

Even if the CET usually exceeds the external tariff of the FTA, it is also usually below that of the Nash.[15] Thus GATT's Article XXIV, which restricts external tariffs on a PTA to be no larger than previous external tariffs on average, should not be a binding constraint on the solution.

3. The static game and the core

The static game consists of two stages: trade agreement choice and tariff setting. This section characterizes the core of the coalitional game under the assumption that countries cannot in the second stage deviate from the tariff agreement made in the first stage. This is a characterizing feature of the static game, and an assumption that will be relaxed in the dynamic setting.

In the first stage, countries are free to make trade agreements. For an agreement to result, all members must accept it.[16] Furthermore, a country accepts an agreement only if it cannot form another agreement (alone or with partner(s)) in which all members in the new agreement yield at least as much welfare (with at least one strict inequality) as under the old agreement. Thus, the resulting trade agreements are exactly those that are in the core, as in Riezman (1985). The advantage of this solution concept is, of course, that all countries behave optimally and that agreements chosen by welfare maximization are more easily enforced than externally determined agreements.

To solve the static game analytically, I proceed in two steps. First, a mapping from trade agreement to welfare is constructed using the second stage

[15] For the XY-CU, this is true provided

$$(b_x + b_y)\,(14a_x - 58c_x + 37c_y + 7c_z) + 100b_x\,(a_x - a_y) > 0.$$

[16] This corresponds to unanimous regionalism in Yi (1996).

optimal tariffs. This is the aim of the next subsection. Second, I let policymakers use this mapping to make trade agreement choices as described above. In what follows, it will be convenient to number the eight trade agreements under consideration:

Table 1. Trade agreements.

1.	Nash equilibrium
2.	FTA between X and Y
3.	FTA between X and Z
4.	FTA between Y and Z
5.	CU between X and Y
6.	CU between X and Z
7.	CU between Y and Z
8.	Free trade

Furthermore, denote the trade agreements γ and the consumption of the numeraire and non-numeraire good in each country under these trade agreements T_γ: $T_\gamma = \{\xi_x, q_x, \xi_y, q_y, \xi_z, q_z\} \mid \gamma$ for $\gamma = 1, \ldots, 8$.

3.1. Welfare revisited

To determine the mapping from trade agreement to welfare for each country, I simply insert the second stage optimal tariffs into the welfare functions (8). The fact that optimal tariffs are unique means that the mapping will be well defined for all countries and agreements in Table 1.

For tractability, I subsequently assume

$$a_x = a_y = a_z = a \quad \text{and} \quad c_x = c_y = c_z = c.$$

This causes optimal tariffs to be independent of any differences in demand between countries:

$$t^I = \frac{3}{10}(a-c), \quad t^{FTA} = \frac{1}{7}(a-c), \quad \text{and} \quad t^{CU} = \frac{5}{19}(a-c)$$

where the interpretation of superscripts should be evident.[17] Internalization of the partner's welfare causes tariffs in a CU to be almost double of those of the FTA, though still more than 10% below tariffs in the Nash. Because

$$t^I > t^{CU} > t^{FTA}$$

[17] Freund (2000), p. 364 finds similar optimal tariffs for symmetric countries.

always, the potential problem of a PTA increasing average external tariffs and thus violating the GATT requirement will not be encountered. Note also that the simplification causes FTA-partners to set the same external tariffs even when they are asymmetric.

Now, let the "size" of demand in country i be defined as

$$d_i = \frac{(a-c)^2}{b_i}.$$

Then the welfare under free trade is:

$$W_x^F = \frac{9}{32}d_x + \frac{1}{16}d_x + \frac{1}{16}d_y + \frac{1}{16}d_z + 0 + 0 \tag{13}$$

where the terms correspond to consumers surplus, producers surplus at home, producers surplus in Y and Z, and tariff revenue from import from Y and Z respectively. In this setting, firms have no advantage in their home market compared to other firms (there are no tariffs) which is why producer surplus follows the same formula at home and abroad.

In the Nash equilibrium, welfare is given by:

$$W_x^I = \frac{9}{50}d_x + \frac{4}{25}d_x + \frac{1}{100}d_y + \frac{1}{100}d_z + \frac{3}{100}d_x + \frac{3}{100}d_x \tag{14}$$

where the interpretation of terms is as above. As in Brander and Spencer (1984), pure profits are shifted from foreign firms to the domestic country in terms of profits for the domestic firm and tariff revenue. The cost, of course, is an anticompetitive effect which causes domestic consumption and thus consumers surplus to fall. Because the two other countries adopt similar beggar-thy-neighbor policies, export profits decrease. Overall, country X is worse off than under free trade unless $45d_x > 42\,(d_y + d_z)$. That is, unless country X has such big demand that gains from rent extraction through tariffs overshadow efficiency gains obtainable from free trade.

The welfare achieved under free trade and the Nash serve as comparison points for welfare obtainable under PTAs. Performing this comparison yields Proposition 3:

Proposition 3. *For all countries and preferential agreements*

- *A country is better off in a PTA than under the Nash provided the market of the partner country is sufficiently large.*
- *A country is better off in a PTA than under free trade provided the joint market of member countries is sufficiently large.*
- *A country prefers to form a PTA with the country that has the largest market of the two other countries.*
- *A country is better off in an FTA than in a CU (with the same partner), provided the country's own demand is at least $\frac{40}{21}$ times the demand of the partner country.*

where what constitutes "sufficiently large" depends on whether the PTA in question is an FTA or a CU. In what follows, I will go through the welfare effects of FTAs and CUs. Proposition 3 follows from the discussion.

The welfare under a PTA between X and Y is given by:

$$W_x^{FTA,XY} = \frac{25}{98}d_x + \frac{4}{49}d_x + \frac{4}{49}d_y + \frac{1}{100}d_z + 0 + \frac{1}{49}d_x \tag{15}$$

$$W_x^{CU,XY} = \frac{169}{722}d_x + \frac{36}{361}d_x + \frac{36}{361}d_y + \frac{1}{100}d_z + 0 + \frac{5}{361}d_x \tag{16}$$

where the interpretation of sub- and superscripts should be obvious. It is seen that overall consumption of the non-numeraire good has decreased and lowered consumers surplus compared to free trade though the level is still above that of the Nash. The preferential treatment of firms operating from within the PTA causes firm X's profits from country Y to increase compared to both the Nash and free trade. This gain can be decomposed into two separate effects: a direct effect on profits caused by the drop in effective cost of production (which falls from $c + t_y^x$ to c), and a strategic effect caused by the trade shift of quantities in favour of firm X because of its new cost advantage.[18]

In the home market, the direct effect of PTA formation is always zero. Compared to the Nash, there is a strategic loss because firm Y now shares firm X's cost advantage: domestic profits fall. Compared to free trade, there is a strategic gain since firm Z loses its tariff-free access to country X: domestic profits increase. Profits from the third country are as under the Nash since this country's tariff against X is independent of the PTA formation. Compared to the Nash, imports from Z rise under an FTA but falls under a CU. Tariff revenue falls in both cases because the revenue from each unit decreases.

For country X to gain from PTA membership compared to the Nash, X must be sufficiently small compared to Y so that preferential access to Y's market outweighs all losses following the opening of X's own market ($210d_x < 351d_y$ for an FTA, $1890d_x < 3239d_y$ for a CU). In the real world, this for instance has the plausible implication that it is more likely that Canada and Mexico gain from NAFTA than that the US does. Finally, country X prefers as partner the country that has the largest demand $\big($i.e. $W_x^{PTA,XY} > W_x^{PTA,XZ}$ iff $d_y > d_z\big)$. This is intuitive as country X is interested in getting preferential access to the largest market possible[19] (due to market segmentation, the cost of gaining that preferential access is the same no matter which partner is chosen).

For a PTA to be preferred over free trade, it must be the case that increased access to the outsider market does not make up for the loss of preferential treatment in the partner country; $175d_x + 250d_y > 686d_z$ for an FTA

[18] Krishna (1998). Obviously, the direct effect is zero when moving from free trade to a PTA.

[19] This is in line with Perroni and Whalley (2000), who argue that the main motive of small countries entering PTAs is to secure their access to a large market. The result more generally helps explain why small countries often make concessions to large countries when entering a PTA.

and $1125d_x + 10750d_y > 15162d_z$ for a CU. A large domestic demand is required because the gains from rent shifting (in form of domestic profits and tariff revenue) are proportional to domestic demand (so is the loss in consumer surplus, but this effect is outweighed). The size of the partner's market is even more important because this determines the demand to which the domestic firm gains preferential access.[20]

The trade off offered between a CU and an FTA is one of producer surplus versus consumer surplus and tariff revenues: the former is larger under the CU (CU markets are more protected) while the latter are greater under an FTA for the same reason. Because the effects on producers in the home market are always outweighed by the effects on consumers and tariff revenues, the size of the partner's market becomes determinant for whether an FTA or a CU is preferred. Formally, domestic demand must be at least $\frac{40}{21}$ times the demand in the partner country for an FTA to be preferred over a CU. However, this is naturally highly sensitive to the relative weights used in the welfare function.

The previous Proposition shows how countries value PTA membership, but in order to be able to compare a country's welfare under all possible agreements in Table 1, it is necessary to also determine the welfare of outsider countries. This leads to the following Proposition:

Proposition 4. *A country in the Nash is better off than outside a CU, but worse off than outside an FTA. Furthermore, a country is best off not making any cooperative agreement (whether or not the two remaining countries form a PTA) provided its own market is sufficiently large.*

The first part of Proposition 4 is easily proved by comparison of outsider-welfare, given by

$$W_x^{FTA,YZ} = \frac{9}{50}d_x + \frac{4}{25}d_x + \frac{1}{49}d_y + \frac{1}{49}d_z + \frac{3}{100}d_x + \frac{3}{100}d_x \tag{17}$$

$$W_x^{CU,YZ} = \frac{9}{50}d_x + \frac{4}{25}d_x + \frac{1}{361}d_y + \frac{1}{361}d_z + \frac{3}{100}d_x + \frac{3}{100}d_x \tag{18}$$

to the welfare under the Nash. Again, the interpretation of sub- and superscripts should be obvious. The reason why a country is better off outside an FTA than in the Nash is that the tariffs levied against it decrease while its domestic market is unaffected by the coalition formation. Thus, exports increase and the outsider country gains. Macho-Stadler et al (1998) find a similar effect in their symmetric model, and describe it as the outsider free riding on the others- a problem which may be severe enough to prevent multilateral agreements.[21] For the CU, the domestic market is still unaffected, but the high CET of the CU causes firm X to earn less abroad so that country

[20] Bond et al (2003) reach a similar result for FTAs, though their intuition is that if the size of FTA-members is large, a larger fraction of trade is internal ex ante, and the FTA suffers less when the rest of the world reacts to its tariff decrease by raising their tariffs.

[21] In Bond et al (2003), the outsider to an FTA gains compared to the Nash because the FTA's reduction in external tariffs causes terms of trade to rise for the outsider.

X is unambiguously worse off than under the Nash. This is true even though $t^I > t^{CU}$.[22] To see why, consider market Y where X's profits are given by $\frac{1}{16b_y}\left(a - c - 3t_y^x + t_y^z\right)^2$. The fall in t_y^x does increase profits abroad, but not enough to outweigh the adverse effect of firm Z gaining preferential access to market Y (t_y^z falls to zero).

To prove the last part of Proposition 4, it is necessary to show that the lowest welfare a country not cooperating with anyone can achieve (i.e. that of a CU-outsider) exceeds welfare under both PTA membership and free trade provided domestic demand is sufficiently large. First, for a CU non-member to be better off than in a CU (FTA) with Y, $d_x > \frac{3500}{1890}d_y + \frac{261}{1890}d_z$ $\left(d_x > \frac{4650}{2527}d_y + \frac{609}{3610}d_z\right)$ is required: X must be so large compared to Y that preferential access to market Y does not sufficiently compensate X giving Y preferential access to its own market. At the same time, Z cannot be too big since this would cause X's lack of preferential access to Z under the YZ-CU to be too costly. Second, it is easily shown by subtraction that X is better off as a CU-outsider than under free trade when $d_x > \frac{1150}{1083}(d_y + d_z)$; country X must be so big that gains from protection with Nash tariffs exceed the combined efficiency gains of free trade and the losses from lack of preferential access under the PTA.

3.2. The core

From the trade agreement-welfare mapping, policy makers determine how a specific agreement affects them. They then maximize welfare by choosing a trade agreement under the constraint that any partners must also accept the agreement for it to be possible. For tractability, I subsequently assume that X and Y are identical in all respects: $d_x = d_y = d$.

The trade agreements under consideration are the eight possibilities listed in Table 1. Formally, the trade agreement γ ($\gamma \in (1, 2, \ldots, 8)$) is in the core iff it is unblocked, where γ is blocked iff there exists a coalition β which yields a consumption $S = \{\xi_x, q_x, \xi_y, q_y, \xi_z, q_z\} \mid \beta$ such that $U_i(S) \geq U_i(T_\gamma)\ \forall\ i \in \beta$ with at least one strict inequality.[23] While the utility of two- and three country coalitions are well defined because the resulting trade agreements are unique, the utility of single country coalitions depends on what the remaining two countries do. I assume that they do what is most likely, namely optimize their welfare given the constraint that the third country will not cooperate. Because of the symmetry between X and Y, I will not encounter the problem that one of the remaining countries prefers one agreement while the other prefers something else.

All countries have preferences over trade agreements given d, d_z. These are determined by comparing the welfare obtained for a country under different

[22] Syropoulos (1999) also finds that the formation of a CU may not benefit the rest of the world even if the CET of the CU falls.

[23] This paper follows Riezman (1985) and Macho-Stadler et al (1998) in their use of core theory in that I do not require the blocking coalition to be better off *on its own* compared to under the blocked agreement: countries belonging to the blocking coalition do trade with countries not in the coalition.

agreements, using the results from the previous subsection. For country X:

Table 2. Preferences over trade agreements.

$\frac{1}{14}d > d_z$	:	$T_5 \succ T_2 \succ T_4 \succ T_1 \succ T_8 \succ T_7 \succ T_3 \succ T_6$
$\frac{37}{110}d > d_z > \frac{1}{14}d$	:	$T_5 \succ T_2 \succ T_4 \succ T_8 \succ T_1 \succ T_7 \succ T_3 \succ T_6$
$\frac{21\,007}{46\,500}d > d_z > \frac{37}{110}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_1 \succ T_7 \succ T_3 \succ T_6$
$\frac{1629}{3500}d > d_z > \frac{21\,007}{46\,500}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_1 \succ T_3 \succ T_7 \succ T_6$
$\frac{21}{40}d > d_z > \frac{1629}{3500}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_1 \succ T_3 \succ T_6 \succ T_7$
$\frac{1890}{3239}d > d_z > \frac{21}{40}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_1 \succ T_6 \succ T_3 \succ T_7$
$\frac{70}{117}d > d_z > \frac{1890}{3239}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_6 \succ T_1 \succ T_3 \succ T_7$
$\frac{425}{686}d > d_z > \frac{70}{117}d$	:	$T_5 \succ T_2 \succ T_8 \succ T_4 \succ T_6 \succ T_3 \succ T_1 \succ T_7$
$\frac{625}{798}d > d_z > \frac{425}{686}d$	:	$T_5 \succ T_8 \succ T_2 \succ T_4 \succ T_6 \succ T_3 \succ T_1 \succ T_7$
$\frac{4827}{6100}d > d_z > \frac{625}{798}d$	:	$T_8 \succ T_5 \succ T_2 \succ T_4 \succ T_6 \succ T_3 \succ T_1 \succ T_7$
$\frac{87}{100}d > d_z > \frac{4827}{6100}d$	:	$T_8 \succ T_5 \succ T_2 \succ T_6 \succ T_4 \succ T_3 \succ T_1 \succ T_7$
$\frac{143\,511}{158\,711}d > d_z > \frac{87}{100}d$	:	$T_8 \succ T_5 \succ T_2 \succ T_6 \succ T_3 \succ T_4 \succ T_1 \succ T_7$
$d > d_z > \frac{143\,511}{158\,711}d$	:	$T_8 \succ T_5 \succ T_6 \succ T_2 \succ T_3 \succ T_4 \succ T_1 \succ T_7$
$\frac{7469}{6669}d > d_z > d$	:	$T_8 \succ T_6 \succ T_5 \succ T_3 \succ T_2 \succ T_4 \succ T_1 \succ T_7$
$\frac{14\,037}{10\,750}d > d_z > \frac{7469}{6669}d$	:	$T_8 \succ T_6 \succ T_3 \succ T_5 \succ T_2 \succ T_4 \succ T_1 \succ T_7$
$\frac{30}{17}d > d_z > \frac{14\,037}{10\,750}d$	:	$T_6 \succ T_8 \succ T_3 \succ T_5 \succ T_2 \succ T_4 \succ T_1 \succ T_7$
$\frac{511}{250}d > d_z > \frac{30}{17}d$	:	$T_6 \succ T_8 \succ T_3 \succ T_5 \succ T_4 \succ T_2 \succ T_1 \succ T_7$
$\frac{2510}{969}d > d_z > \frac{511}{250}d$	:	$T_6 \succ T_3 \succ T_8 \succ T_5 \succ T_4 \succ T_2 \succ T_1 \succ T_7$
$d_z > \frac{2510}{969}d$	:	$T_6 \succ T_3 \succ T_8 \succ T_4 \succ T_5 \succ T_2 \succ T_1 \succ T_7$

To understand these rankings, consider for instance agreement 6, the CU between X and Z. In the beginning of the table, demand is very low in Z compared to X and Y, and so a CU with Z is the worst outcome for X: X looses tariff revenue from Z and has to open its market but gets only preferential access to a very small country in return. However, as one moves down the table, the relative demand in Z increases, and it becomes much more desirable for X to have an XZ-CU. In fact for $d_z > \frac{14\,037}{10\,750}d$, this is X's best option as it simultaneously gives the country preferential access to a large market (as opposed to the Nash or any trade agreement with Y) and lets it collect tariffs from Y (as opposed to free trade). Of course, the opposite holds for country Z,

and so this CU cannot necessarily be optimal for X and Z at the same time. In fact, when both X and Y are large compared to Z, it is optimal for Z to have free trade as this secures free access to two large markets where the XZ-CU would only give preferential access to one. Thus, Z never prefers an XZ-CU over all other alternatives. However, as we will see below, this CU still belongs to the core under certain conditions.

Free trade starts out in the middle of the ranking in Table 2, and becomes increasingly desired as countries get closer in size. It is the preferred outcome when countries are fairly similar, but then decreases in desirability as preferential access to Z becomes more profitable in terms of welfare. All other trade agreements move up and/or down the preference scale of country X for similar reasons. Country Y's ranking of the alternatives is symmetric to the one of X, and a similar, though not symmetric, ranking can be constructed for country Z.

The trade agreements in the core can now be determined by simply comparing the rankings of all countries and see which agreements are unblocked. Before stating the actual result, it is worthwhile to consider some special cases to understand the outcome. For $\frac{625}{798}d > d_z$, the XY-CU is preferred by both X and Y and will therefore block any other trade agreement, irrespective of the fact that this is the worst possible allocation for Z. Thus, in a world with two large countries and one small, one would expect the two large countries to collude into a CU. When the three countries are more similar in size $\big($that is for $\frac{14\,037}{10\,750}d > d_z > \frac{625}{798}d\big)$, free trade is preferred by everyone and is therefore the only agreement in the core.

When d_z increases further compared to d, agreements in the core are not necessarily the preferred outcome for all members of the agreement. Consider the interval $\frac{121\,889}{92\,610}d > d_z > \frac{14\,037}{10\,750}d$, for which preferences are given by:

$$\begin{aligned} X\colon\ & T_6 \succ T_8 \succ T_3 \succ T_5 \succ T_2 \succ T_4 \succ T_1 \succ T_7 \\ Y\colon\ & T_7 \succ T_8 \succ T_4 \succ T_5 \succ T_2 \succ T_3 \succ T_1 \succ T_6 \\ Z\colon\ & T_8 \succ T_7 = T_6 \succ T_2 \succ T_4 = T_3 \succ T_1 \succ T_5 \end{aligned} \tag{19}$$

where $T_7 = T_6$ and $T_4 = T_3$ for Z follows from the symmetry of X and Y. Clearly the Nash, the XY-FTA, the XY-CU, and the XZ- and YZ-FTAs are all blocked by free trade. Both X and Y would like to make a CU with Z, and even if Z prefers free trade to this, it cannot block it as not all members of the free trade coalition would obtain as much welfare as under the CU. Similarly for free trade: it is preferred by Z, and neither X nor Y alone or together can block this. Thus, the core contains both the XZ- and YZ-CUs and free trade.

For all remaining values of relative demand, the core is found in a similar manner. This results in Proposition 5:

Proposition 5. *Given demand d, d_z, the trade agreements most likely to result*

(i.e. the agreements in the core) are:

$$\begin{aligned}
&\textit{For } \tfrac{625}{798}d > d_z && : XY\textit{-CU}\\
&\textit{For } \tfrac{14\,037}{10\,750}d > d_z > \tfrac{625}{798}d && : \textit{Free trade}\\
&\textit{For } \tfrac{3761}{1890}d > d_z > \tfrac{14\,037}{10\,750}d && : \textit{Free trade, } XZ\textit{-CU, } YZ\textit{-CU}\\
&\textit{For } \tfrac{2300}{1083}d > d_z > \tfrac{3761}{1890}d && : \textit{Free trade}\\
&\textit{For } d_z > \tfrac{2300}{1083}d && : XY\textit{-CU}
\end{aligned}$$

This result is interesting for several reasons. First of all, the solution concept sharpens predictions from the eight possible trade agreements in Table 1 to just four outcomes. Furthermore, the core is non-empty so that it is possible to make a prediction for any relative demand, and the prediction will be precise unless dealing with two small and one slightly larger country. Interestingly, the model predicts that a cooperative agreement will be reached irrespective of relative demand.

It is seen that efficiency gains from free trade unambiguously outweigh the preferential effects of a PTA when countries are rather similar. This is in contrast to Riezman (1985) who in an example shows that all possible (2-country) CUs can be in the core when countries are symmetric.

In the third case in Proposition 5, the fact that XZ-, YZ-CUs are possible outcomes even when free trade is actually preferred by Z shows that even large countries can lose compared to free trade by *voluntarily* entering a CU agreement. It also shows that a country can lose compared to a CU by entering a free trade agreement just as voluntarily.[24] These observations lead to the following Proposition:

Proposition 6. *In the static game, it is possible for a country to lose compared to free trade by voluntarily entering a CU. Furthermore, countries may lose compared to a CU by voluntarily agreeing to free trade.*

The first part of Proposition 6 is strengthened by observing that even if the XY-CU is the predicted outcome both for very small and very large relative demand, the reasons for the two predictions are very different. In particular, when markets in X, Y are very small compared to Z, Z prefers to act alone no matter what X and Y do (cf. Proposition 4). Due to Z's lack of cooperation, X and Y are forced to settle with a CU between them even if it is only their 4^{th}-5^{th} best option. As in Riezman (1985) example 1, the small countries *do not gain from the CU compared to free trade*, but rather choose the CU as a best response to a country that gains from a tariff war. Also, in this setting, large countries will never support free trade unless sufficiently compensated to

[24] Given any of the three outcomes, countries act voluntarily in the sense that they cannot force a different agreement which makes them better off.

do so by the small country(ies). On the other hand, when relative demand in Z is very small, X and Y are not interested in sharing their markets with Z. Instead, they form an XY-CU to have mutual access to X, Y markets while still being able to collect tariff revenue from Z.

It is striking that the core contains no FTAs at all. The analytical argument for this is that for two countries to rank an FTA over a CU, demands must be so different that the large country involved will either be better off on its own (in Z's case) or in a CU with a country of equal size (in X's case). In practice, the creation of an FTA can also undermine the individual countries' bargaining power, and the administrative cost of rules of origin is very high.

However, this immediately raises the question of why most PTAs in the world are then FTAs and not CUs? One explanation is political. For instance, the facts that negotiations for an FTA can be easier than those for a CU because CETs need not be agreed upon and that FTAs provide greater flexibility in trade policy against the rest of the world are not captured by this model.[25] Furthermore, it is possible that (as with the Andean Community) FTAs are only a first step towards deeper integration, and that many FTAs will turn into CUs over time.

4. The dynamic game

In the static game, trade agreements are predicted to be coalition formations in the core. The dynamic game builds on a similar intuition, but goes further in that it restricts the trade agreements available to countries to those that are in fact sustainable against deviation (to be defined).

By imposing this requirement, the section makes the model more realistic in two ways: countries have the opportunity to cheat, and the game is modelled over several periods. As the example in the introduction shows, an incentive to cooperate does not always equal the ability to do so. Thus, one should model endogenously formed agreements as self-sustaining.

I show below that none of the solutions in the core are robust to deviation when only one period is considered. In fact, of all options available to countries in Table 1, only the Nash is robust because it is non-cooperative to begin with. The subsequent inclusion of more periods imposes a cost on cheating which in some circumstances is sufficient to restore bilateral and/or multilateral cooperation. Furthermore, some results from the static case do carry over to the dynamic framework provided countries are patient.

4.1. Sustainability

In the present paper, sustainability is defined in a standard repeated game framework by invoking the Folk Theorem: an agreement is sustainable iff the

[25] The latter could be included in the model by allowing a to differ between countries in the analysis of the core.

welfare obtained by adhering to the agreement exceeds that obtainable by deviating in one period and then being punished by moving to an alternative continuation equilibrium that yields a lower average payoff than the original agreement. Countries can only deviate individually, and I assume that when they do, their punishment consists of return to the Nash forever. While this particular punishment strategy is chosen for tractability, it should be noted that there are many different alternatives which may sustain different outcomes.

With this definition of punishment, any agreement which yields at least as much welfare as the Nash for all members can be sustained provided countries are sufficiently patient. This subsection determines exactly how patient countries need to be for a given trade agreement to be sustainable given relative demand. The selection between sustainable equilibria is discussed in the next subsection.

To formally define sustainability, it must first be defined how a deviation takes place, who is able to deviate, and what they gain from doing so. To deviate, a country agrees to some coalition as usual in the first stage of the coalition formation game. But when it comes to setting tariffs in the second stage, the country sets Nash tariffs instead of the tariffs it would set if maximizing utility subject to the constraints of the trade agreement (remember that the optimal tariffs of a country not bound by any agreement are the Nash tariffs always). Tariffs are set simultaneously, so the other countries cannot react to the deviation in the period in which it occurs: they set their tariffs according to the stage one agreement. The latter is what defines the difference between a deviation and a single country blocking since blocking is only possible if welfare increases for the blocking country *once outsider countries have adjusted their tariffs.*[26]

Any country can deviate from an agreement, but if a non-member does so it has no effect on member cooperation (for the outsider, deviation and compliance is the same thing since Nash tariffs are played in both cases). Deviation by a member causes that member to gain because at least one binding constraint is removed from its maximization problem while everything else (that is, the tariffs of the other countries) is kept constant. The deviation gain is defined as the difference between the welfare obtained by deviation and that obtained by adherence. Using country X as an example, the former is defined as:

$$W_x^{D\gamma} = W_x\left(t_x^y(1), t_x^z(1), t_y^x(\gamma), t_y^z(\gamma), t_z^x(\gamma), t_z^y(\gamma)\right) \tag{20}$$

where $W_x^{D\gamma}$ is country X's welfare when deviating from γ, and the numbers/γ in the parentheses stand for the trade agreement according to which that tariff is set.

It can thus already be observed that no trade agreements (except the non-cooperative one) is robust to such deviation in a one period game. Therefore, one would expect to observe Nash behavior only. Naturally, this is not the case,

[26] One may think of a single country blocking as taking place in the first stage where trade agreements are formed so that all other countries adjust their tariffs optimally, whereas a deviation comes as a surprise in the tariff setting stage so that the other countries cannot react.

and subsequent periods must be included in the analysis. I do so by assuming that countries play trigger strategies such that a defection will be punished by playing Nash tariffs forever.[27] This punishment is credible because Nash tariffs are indeed optimal in the non-cooperative game. When this threat is large enough to influence current behaviour towards cooperation, the trade agreement is sustainable:

Definition 7. *A trade agreement γ is sustainable against deviation iff*

$$W_i^{D\gamma} + \sum_{\tau=1}^{\infty} \delta^\tau W_i^I \leq \sum_{\tau=0}^{\infty} \delta^\tau W_i^\gamma \quad \forall i \in \gamma$$

$$\iff \quad \left(W_i^{D\gamma} - W_i^\gamma\right) \leq \frac{\delta}{1-\delta}\left(W_i^\gamma - W_i^I\right) \quad \forall i \in \gamma$$

where δ is the discount factor which is assumed common for all countries ($\delta \in [0,1)$), W_i^γ is the welfare obtained by country i under trade agreement γ, and $i \in \gamma$ iff i is a member of γ. The first line shows that an agreement is sustainable iff the welfare in the deviation period plus the discounted welfare in all punishment periods is less than the discounted welfare in the trade agreement. The second line compares the gains and losses relative to the welfare obtained under γ: γ is sustainable if and only if the deviation gain is no greater than the discounted punishment loss.

By insertion of the relevant tariffs into the welfare function and solution of the above inequalities, I obtain the requirement for sustainability for all possible agreements. (See Table 3.)

First of all, for any country to be willing to participate in a cooperative trade agreement, it must be the case that the welfare obtained by that country under the cooperative agreement exceeds that achievable under the Nash. If not, the deviating country would not only gain in the deviation period, but also in all subsequent "punishment" periods and would therefore definitely deviate. This is what gives rise to the requirements on demand in Table 3.[28] Trade agreements 2 and 5 do not contain such requirements because X and Y are symmetric, and the welfare from a CU or FTA between symmetric countries always dominates that obtainable in the Nash. The symmetry between X and Y is also what causes the sustainability requirements for agreements 3-4 and 6-7 to be identical.

Given the restrictions on demand are satisfied, the deviating country will gain in the deviation period and lose in every punishment period compared to staying in the original agreement. Whether deviation pays or not therefore comes to depend on the deviating country's degree of patience, which is measured by δ. As always, sustainable cooperation is increasing in this parameter.

[27] It is easily shown that the trigger strategy is optimal when other players play the same strategy, see Gibbons (1992).

[28] Equivalently, $\delta \geq 1$ must hold outside these intervals for an agreement to be sustainable. This is impossible since $\delta \in [0,1)$.

Table 3. Requirement for sustainability.

	Agreement sustainable provided	
T_1	–	–
T_2	–	$\delta \geq \frac{70}{117}$
T_3	$\frac{117}{70}d > d_z > \frac{70}{117}d$	$\delta \geq \max\left\{\frac{70}{117}\frac{d}{d_z}, \frac{70}{117}\frac{d_z}{d}\right\}$
T_4	$\frac{117}{70}d > d_z > \frac{70}{117}d$	$\delta \geq \max\left\{\frac{70}{117}\frac{d}{d_z}, \frac{70}{117}\frac{d_z}{d}\right\}$
T_5	–	$\delta \geq \frac{1890}{3239}$
T_6	$\frac{3239}{1890}d > d_z > \frac{1890}{3239}d$	$\delta \geq \max\left\{\frac{1890}{3239}\frac{d_z}{d}, \frac{1890}{3239}\frac{d}{d_z}\right\}$
T_7	$\frac{3239}{1890}d > d_z > \frac{1890}{3239}d$	$\delta \geq \max\left\{\frac{1890}{3239}\frac{d_z}{d}, \frac{1890}{3239}\frac{d}{d_z}\right\}$
T_8	$\frac{28}{15}d > d_z > \frac{1}{14}d$	$\delta \geq \max\left\{\frac{15}{14}\frac{d}{d_z+d}, \frac{15}{14}\frac{d_z}{2d}\right\}$

All threshold values for δ depend on the relative demand in member countries only, since this is what determines the relative sizes of gains and losses connected with the agreement. In particular, demand in the non-member country is irrelevant because members' sales in this country are the same under compliance with the agreement, deviation from it, and Nash punishment. Thus, all expressions containing non-member demand cancel from the second line in Definition (7) above. It follows that the threshold of δ is independent of demand when member countries are symmetric.

To have an overview of how relative demand otherwise affects the threshold of δ, consider Figure 2. In this figure, the horizontal axis measures relative demand $\left(\frac{d_z}{d}\right)$ while the vertical axis measures delta ($\delta \in [0, 1)$). FTAs between asymmetric countries are sustainable above the line EIF, CUs above the line DJG, and free trade sustainable above AKH. An FTA between symmetric countries is sustainable above BL, and a CU sustainable above CM. It is clear that PTAs are more difficult to achieve when partners are asymmetric, because both members' incentive to deviate must be taken into account: asymmetry causes one partner to have a greater interest in the agreement (because access to a larger market is gained), while the other partner is less interested. Because the most strict condition is determinant for the overall requirement for sustainability, this causes an increase in the threshold on δ. Furthermore, the threshold is increasing in the degree of asymmetry. The same result holds for free trade sustainability.

The reason why CUs are easier to sustain than FTAs is that the punishment loss is greater (at least for the relevant values of relative demand). This again follows from the internalization of effects on the partner's welfare in a CU when

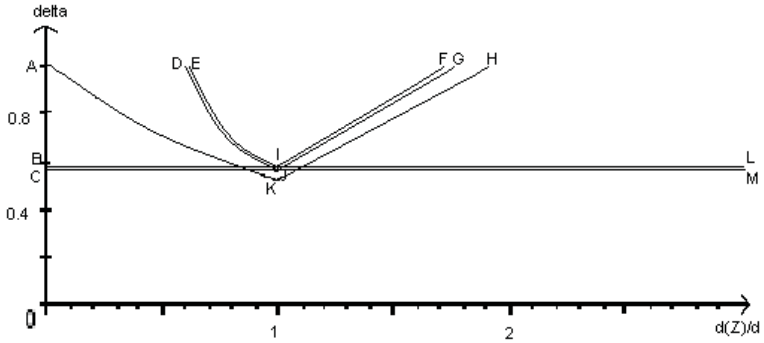


Figure 2. The dependence of sustainability on relative demand.

setting the CET. The gain in the deviation period also increases, though not enough to outweigh the effect of increased punishment. Free trade is even easier to sustain than a CU, again because an increased punishment loss outweighs the effect of an increase in the deviation gain.

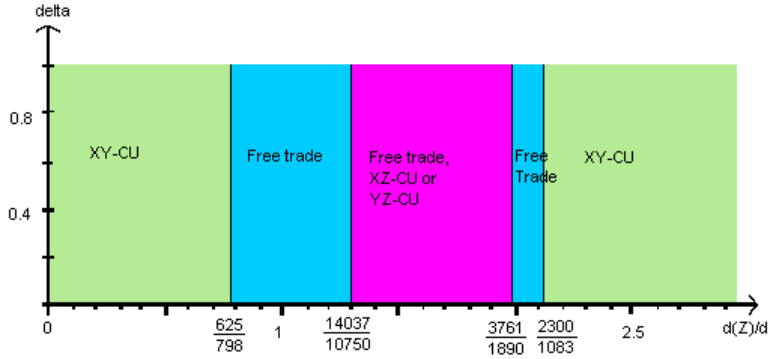
4.2. The game

The dynamic game consists of infinitely many periods with one tariff-setting stage in each period. Before the game starts, there is a stage "0" where countries coordinate to a certain equilibrium according to the core, where the core is the outcome of a hypothetical initial bargaining process where blockings and counterblockings are also hypothetical. That is, the core is used as an equilibrium selection criterion which finds the most likely outcome of the bargaining process among all sustainable trade agreements.

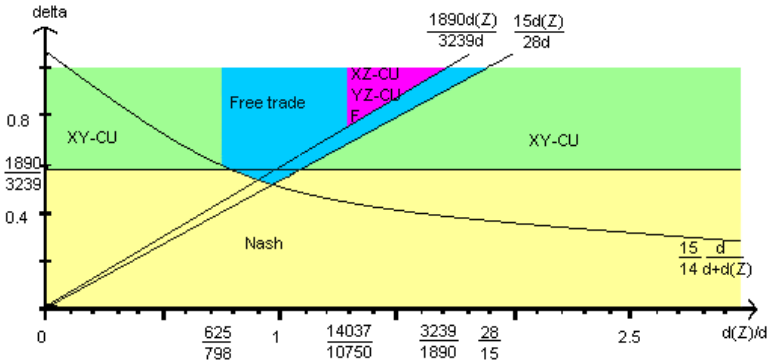
The outcome of all previous periods is observed by all countries before the current period, and countries play the previously described trigger strategy. I assume complete information so that each country knows whether it would pay for its partners to deviate. Because subsequent optimal behaviour is taken into account at stage "0", this means that an unsustainable trade agreement will never be made to begin with. Thus, the trade agreement signed in stage "0" will continue to hold forever in this setup.

In practice, the game is solved as follows: for all possible combinations of relative demand and delta, I determine which agreements are sustainable, and use *these agreements only* as available possibilities in a hypothetical static game, which is solved as the static game above by finding the agreements in the core. In this way, I determine the agreements most likely to form out of all sustainable agreements.

The results are shown graphically in Figure 3B and algebraically in the appendix. Figure 3A shows the trade agreements in the core of the static game for comparison. As in figure two, the axes measure relative demand and delta respectively.



A. Solution to the static game



B. Solution to the dynamic game

Figure 3

There are seen to be big differences in the two solutions. Most obviously, Figure 3B contains a large area where the solution is non-cooperative because one (or more) country(ies) is (are) too impatient to maintain any cooperative trade agreement: the short term gain of a deviation is preferred to the infinite stream of gains from cooperation. The fact that cooperation is reduced mainly for low discount factors can be problematic for real world cooperation since governments sit for a short period only, and therefore often have low discount factors.[29]

The interval of relative demands for which an XY-CU can be the solution has increased. This expansion, however, comes at the expense of a reduction in the area for which free trade can be a solution, and occurs because Z will find it optimal to deviate from free trade even if its relative demand is smaller than before. There are two reasons for this, and which is applicable depends very much on the size of relative demand.

[29] Dixit (1987).

When $\frac{d_z}{d} < \frac{28}{15}$, Z is better off under free trade than in the Nash, but for some values of relative demand not more so than that it pays for Z to deviate should a free trade agreement form. X and Y realize this and choose their best possible sustainable option, the XY-CU. When $\frac{2300}{1083} > \frac{d_z}{d} > \frac{28}{15}$, the reason has no longer much to do with sustainability, but rather comes from the assumption about how deviation is punished and the fact that for Z, the Nash is preferred to free trade which again is preferred to any CU for this interval of relative demand. While in the static case, Z would not block free trade because the resulting XY-CU would make it worse off, the possibility of a deviation not only allows Z to collect a gain in the deviation period, but also to gain from being in the Nash instead of in free trade in all subsequent periods. Thus, Z will definitely deviate and free trade becomes infeasible. Again, X and Y choose their best feasible option, the XY-CU.

The comparison of Figures 3A and B gives rise to a central Proposition of this paper:

Proposition 8. *The introduction of the self-sustainability requirement reduces the overall scope for cooperative trade agreements. Free trade is among the most likely outcomes for less values of relative demand while preferential trade agreements are among the most likely outcomes for more values of relative demand. This is true even when countries are very patient (i.e. δ almost equal to one).*

Proposition 8 gives a more negative view on the scope for cooperation compared to the static solution. There is, however, also good news. Most importantly, it has been shown that for the values of relative demand and impatience at which free trade or a PTA is the solution in the dynamic setting, there exists at least one strategy that can support these as equilibrium outcomes when both deviations and blockings are allowed for. Furthermore, when $\frac{28}{15}d > d_z > \frac{14\,037}{10\,750}d$, the sustainability requirement makes free trade the most likely outcome for some values of δ whereas the core of the static game contained both free trade, XZ-CUs, and YZ-CUs. Thus, the requirement facilitates multilateral cooperation in this case. Interestingly, for $\frac{3239}{1890}d > d_z > \frac{14\,037}{10\,750}d$, impatience can cause free trade to be the outcome of the dynamic game, where more patience may yield a PTA instead $\big($since for $\delta \geq \frac{1890}{3239}\frac{d_z}{d}$, there are three possible solutions, but for $\frac{1890}{3239}\frac{d_z}{d} > \delta \geq \frac{15}{28}\frac{d_z}{d}$ free trade is the only solution$\big)$.

The conclusion that FTAs are not among the most likely equilibria for any value of relative demand continues to hold. So do the facts that it is possible for a country to lose compared to free trade by voluntarily entering a CU as well as lose compared to a CU by voluntarily agreeing to free trade.

While these results are interesting in their own right, they also raise some questions. In particular, I have not allowed for different punishment strategies or for the non-deviating countries to cooperate during the punishment phases. The implementation of the latter would, for instance, be expected to alter the conclusion that Z will deviate from free trade when $\frac{2300}{1083} > \frac{d_z}{d} > \frac{28}{15}$ (at least for some δs). Dixit [1987] calls the fact that other outcomes may be sustained by

different strategies a non-uniqueness in addition to the one that several different trade agreements may be sustainable by the same strategy. In these terms, the equilibrium selection criteria of this paper solves the latter non-uniqueness (at least for most values of δ and relative demand), but not the former.

4.3. Stepping stone or stumbling block?

The question of whether regional integration will keep economies on the track towards free trade or instead allow regions to turn so powerful that cooperation between them becomes impossible has been a main concern since the beginning of the 1990's when PTAs started to mushroom. Since most models find free trade desirable, the concern is real and has the worst case scenario of three strong trading blocks (the FTAA, the EU and an Asian block) fighting each other instead of cooperating towards a more efficient outcome.[30]

With a slight abuse of the dynamics inherent in the stepping stone/ stumbling block idea, the question can be answered in the current model. Thus, I say that a PTA is a stumbling block iff a PTA is the expected outcome in the previous dynamic game, while free trade would result absent the possibility of regional cooperation. Also, a PTA is a stepping stone when free trade remains the solution to the dynamic game even if countries are free to form PTAs. With these definitions, the following Proposition holds:

Proposition 9. *Whether a CU can be a stepping stone or a stumbling block to free trade in the dynamic game depends on the relative demand between countries as well as the degree of impatience.*

A recalculation of the dynamic solution when PTAs are not allowed for shows that PTAs can be stumbling blocks to free trade. For instance, this is the case when $\frac{625}{798}d > d_z > \frac{1}{14}d$ and $\delta \geq \frac{15}{14}\frac{d}{d_z+d}$. Here, relative demand in X and Y is so big that they prefer an XY-CU to free trade; the increased access to Z's market following full trade liberalization would not generate sufficient welfare gains for X and Y to cover the losses they would suffer from giving up their preferential access under the CU. But absent the opportunity of forming such CU, X and Y would be better off granting Z access to their markets than by playing the Nash, and free trade would result.

On the other hand, a CU can be a stepping stone when $d > d_z > \frac{625}{798}d$, since free trade is in the core no matter whether PTAs can be formed freely or not in this case. This confirms that a PTA can be a stepping stone towards free trade under the unanimous regionalism rule of Yi (1996).[31]

Letting the definition of stepping stone and stumbling block in the static game be as for the dynamic game, only with "dynamic" exchanged for "static", CUs can be both stepping stones and stumbling blocks in the static game as well. For instance, the XY-CU is a stumbling block when $\frac{625}{798}d > d_z > \frac{1}{14}d$

[30] Krugman (1991) finds that world welfare is minimized at three trading blocks.

[31] Yi finds that a PTA can be both a stepping stone and a stumbling block under this rule, but that the latter is typical.

and a stepping stone when $\frac{14037}{10750}d > d_z > \frac{625}{798}d$ for reasons similar to those of the dynamic case.

4.4. The symmetric case and multilateral tariffs

Because one effect of GATT has been to reduce tariffs multilaterally, this section considers the addition of a multilateral tariff to the set of possible cooperation choices. Thus, countries have the options of free trading, being in the Nash, making a preferential agreement, and reducing tariffs multilaterally. For tractability, I shall assume that all countries are symmetric. I will briefly review previous results for the symmetric case before proceeding to the introduction of the multilateral tariff and the demonstration of how it can improve welfare.

Under symmetry, the model simplifies greatly in that all members of an agreement have the same incentive to stay in or deviate from the trade agreement, and all possible FTAs and CUs are symmetric. In this case, welfare reduces to

$$W^I = \frac{21}{50}d, \quad W^{FTA} = \frac{2199}{4900}d, \quad W^{CU} = \frac{869}{1900}d \quad \text{and} \quad W^F = \frac{15}{32}d,$$

where

$$d = \frac{(a-c)^2}{b}.$$

This allows me to rank welfare unambiguously as

$$W^F > W^{CU} > W^{FTA} > W^I,$$

implying that the efficiency of free trade outweighs the redistributional gains of a PTA irrespective of country size and production costs. Thus, under the core solution concept, one should never observe anything but free trade.

If allowing for deviation, it is known from Table 3 and the fact that $d = d_z$ under the symmetry assumption that free trade is only sustainable provided $\delta \geq \frac{15}{28}$. Sustainability of a CU and FTA requires $\delta \geq \frac{1890}{3239}$ and $\delta \geq \frac{70}{117}$ respectively. Since free trade is easiest to sustain and yields the most welfare, it will be the outcome whenever δ exceeds $\frac{15}{28}$. If δ is below this threshold, the Nash results.

Now, a multilateral tariff is simply a tariff that is common for all countries. By insertion of this common tariff, t, into the welfare function, and letting W^M denote the welfare of each country under t, I find that

$$W^M = \frac{1}{32}\frac{(3(a-c)-2t)(5(a-c)+2t)}{b}.$$

This is decreasing in the common tariff,

$$\frac{d}{dt}\left(W^M\right) = -\frac{1}{8}\frac{a-c+2t}{b} < 0,$$

implying that efficiency gains cause free trade to be preferred to an above zero common tariff always $\left(\text{in fact, } W^F - W^M = \frac{1}{8}\frac{a-c+t}{b}t\right)$. This is so because countries are symmetric; the increase in domestic profits and tariff revenues due to rent shifting are never sufficiently large to outweigh the negative effects on consumers surplus and profits from abroad in this case. Thus, one will still expect to observe free trade only under the core solution concept.

Turning to the dynamic case, it is clear from the welfare comparison that free trade will be chosen whenever it is sustainable. But as it turns out, a multilateral tariff reduction can improve the welfare of all countries when free trade is not sustainable. To see this, consider first the requirement for sustainability of the multilateral tariff $\left(\text{letting } W_x^{DM} \text{ denote the welfare of country } X \text{ when deviating from the multilateral tariff}\right)$:

$$W_x^{DM} + \sum_{\tau=1}^{\infty} \delta^{\tau} W_x^I \leq \sum_{\tau=0}^{\infty} \delta^{\tau} W_x^M \tag{21}$$

which, provided the common external tariff is below that of the Nash $\left(t < \frac{3}{10}(a-c)\right)$, reduces to:

$$\frac{5}{4}\,\frac{3(a-c)-10t}{7(a-c)-10t} \leq \delta \tag{22}$$

This lower bound on delta is below that of free trade whenever $t > 0$. Thus, a multilateral tariff can be sustainable even when free trade is not. By comparing welfare at the Nash and an arbitrary common tariff, it is easily established that the multilateral tariff is preferred to the Nash provided $t < \frac{3}{10}(a-c)$. Because welfare is decreasing in the common tariff, countries can maximize welfare by choosing the lowest possible tariff sustainable. That is, the tariff which fulfills $\delta = \frac{5}{4}\,\frac{3(a-c)-10t}{7(a-c)-10t}$. Any higher tariff reduces welfare, and any lower tariff causes deviation and return to the Nash.

5. Concluding remarks

Though the literature on (preferential) trade agreements has developed in the direction of stability issues and endogenous trade agreement formation, there still seems to be a presumption that at least some (part of an) agreement is binding or exogenously specified.

In a three-country oligopolistic trade model, this paper predicts which trade agreements form when choices of countries are endogenous. In particular, the resulting agreement must be self-enforcing, and no other sustainable agreement (among the possibilities listed in Table 1) in which all members achieve at least as much (with one achieving strictly more) welfare as under the chosen agreement must be able to form. The solution method shows how the core can be used as an equilibrium selection tool to determine the agreement most likely to form among all sustainable agreements.

When countries enter a cooperative trade agreement, they gain from the enhanced access to partner market(s) and increased competition at home, but also suffer losses due to the decrease in tariff revenue and domestic monopoly power. The relative size of these effects, and thus the desirability of the trade agreement, is determined by relative demand. In general, the greater the partner market(s), the more is gained, and the greater the domestic market, the more is lost. Sustainability is defined in a standard repeated game framework.

I find that the introduction of the self-enforcement requirement reduces the overall scope for cooperative trade agreements; allowing for deviation alone makes agreements other than the Nash impossible, and though the introduction of more periods makes cooperation attainable in some cases, the total effect is negative. In the static coalition formation game, only CUs and free trade are among the predicted outcomes with the latter belonging to the core only when countries are fairly similar. In the dynamic case, the Nash is the predicted outcome irrespective of relative demand whenever countries are impatient. Furthermore, free trade is among the most likely outcomes for less values of relative demand while preferential trade agreements are among the most likely outcomes for more values of relative demand.

Some results from the static game are robust to the inclusion of the sustainability restriction. Most importantly, with or without the constraint, FTAs are never among the predicted outcomes, and it is possible for a country to lose compared to free trade by voluntarily entering a CU as well as to lose compared to a CU by voluntarily agreeing to free trade. Also, CUs can be both stumbling blocks and stepping stones to free trade in both cases. A sustainable multilateral tariff agreement between symmetric countries can improve welfare only in the dynamic case and only when free trade itself is not sustainable.

The model is one of intraindustry trade between three countries only, and one should therefore interpret its specific predictions with caution. However, the main point of the paper, namely that self-sustainability is important in both preferential and multilateral cooperation and that the inclusion of this constraint matters, seems more robust. In particular, one would expect that the introduction of a self-sustainability requirement for *all* trade agreements in other models (for instance those mentioned previously) would generally cause a change in the pattern of cooperation otherwise found, with all cooperation breaking down for some degree of impatience. The example of the Andean Community shows how cooperation may be optimal, yet not achievable unless all members are ready to give up short term gains.

While the paper provides a framework for predicting trade agreement outcomes which are both stable and sustainable, several extensions remain. Most importantly the introduction of more cooperative agreements in the game (especially to allow for multilateral tariff cooperation which has already been shown to be the solution to the dynamic game in some cases), consideration of different punishment strategies (dynamic punishment and/or punishment modelled according to WTO rules), and a further development of the dynamics in the model. Additional asymmetry is also desirable, particularly in either the

demand parameter a or the cost parameters c. Such extension would cause a new asymmetry in tariffs, which for instance would allow countries to benefit from the flexibility in external tariff setting under an FTA. Moving beyond the current model, the inclusion of trade specific capital which builds up during PTA membership could cause history to become determinant of the pattern of PTAs as the cost of deviation rises.

References

Bagwell, K., Staiger, R. (1997a). "Multilateral cooperation during the formation of free trade areas". *International Economic Review* 38, 291-319.

Bagwell, K., Staiger, R. (1997b). "Multilateral cooperation during the formation of customs unions". *Journal of International Economics* 42, 91-123.

Baldwin, R. (1995). A domino theory of regionalism. In Baldwin, R., Haaparanta, P., Kiander, J. (Eds.), *Expanding Membership of the European Union*. Cambridge University Press.

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

Bond, E., Riezman, R., Syropoulos, C. (2003). "A strategic and welfare theoretic analysis of free trade areas". *Journal of International Economics*, forthcoming.

Brander, J., Spencer, B. (1984). Tariff protection and imperfect competition. In Kierzkowski, H. (Ed.), *Monopolistic Competition and International Trade*. Clarendon Press: Oxford.

Dixit, A. (1987). Strategic aspects of trade policy. In Bewley, T. (Ed.), *Advances in Economic Theory Fifth World Congress*. Cambridge University Press.

Freund, C. (2000). "Multilateralism and the endogenous formation of preferential trade agreements". *Journal of International Economics* 52, 359-376.

Gibbons, R. (1992). *A Primer in Game Theory*. Harvester Wheatsheaf.

Kemp, M., Wan, J. (1976). "An elementary proposition concerning the formation of customs unions". *Journal of International Economics* 6(1), 95-97.

Kowalczyk, C., Sjostrom, T. (1994). "Bringing GATT into the core". *Economica* 61, 301-317.

Krishna, P. (1998). "Regionalism and multilateralism; a political economy approach". *Quarterly Journal of Economics* 113, 227- 251.

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

Ludema, R. (1994). *On the Value of Preferential Trade Agreements in Multilateral Negotiations*. Mimeo: University of Western Ontario.

Macho-Stadler, I., Perez-Castrillo, D., Ponsati, C. (1998). "Stable multilateral trade agreements". *Economica* 65, 161-177.

Perroni, C., Whalley, J. (2000). "The new regionalism: trade liberalization or insurance?". *The Canadian Journal of Economics* 33, 1-24.

Panagariya, A., Krishna, P. (2002). "On necessarily welfare-enhancing free trade areas". *Journal of International Economics* 57(2), 353- 367.

Riezman, R. (1985). "Customs unions and the core". *Journal of International Economics* 19, 355-365.

Syropoulos, C. (1999). "Customs unions and comparative advantage". *Oxford Economic Papers* 51, 239-266.

Viner, J. (1950). *The Customs Unions Issue.* Carnegie Endowment for International Peace: New York.

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

Webpages (last visited 26.07.2005);

www.warwick.ac.uk/~vxscbl/regionalism.pdf

www.wto.org/english/tratop_e/region_e/regfac_e.htm#top

www.ictsd.org/dlogue/1999-02-01/overviewAndean.pdf

www.mercosurconsulting.net/Articles/article2.html

Appendix: the solution of the dynamic game

Table 4. Algebraic solution for the dynamic game

Requirement on d, d_z	Requirement on δ	Trade agreement
$\frac{625}{798}d > d_z$	$\delta \geq \frac{1890}{3239}$	T_5
$\frac{1475}{1764}d > d_z$	$\frac{1890}{3239} > \delta$	T_1
$\frac{1475}{1764}d > d_z > \frac{625}{798}d$	$\frac{15}{14}\frac{d}{d_z+d} > \delta \geq \frac{1890}{3239}$	T_5
$d > d_z > \frac{625}{798}d$	$\delta \geq \frac{15}{14}\frac{d}{d_z+d}$	T_8
$d > d_z > \frac{1475}{1764}d$	$\frac{15}{14}\frac{d}{d_z+d} > \delta$	T_1
$\frac{14\,037}{10\,750}d > d_z > d$	$\delta \geq \frac{15}{28}\frac{d_z}{d}$	T_8
$\frac{3528}{3239}d > d_z > d$	$\frac{15}{28}\frac{d_z}{d} > \delta$	T_1
$d_z > \frac{3528}{3239}d$	$\frac{15}{28}\frac{d_z}{d} > \delta \geq \frac{1890}{3239}$	T_5
$\frac{28}{15}d > d_z > \frac{14\,037}{10\,750}d$	$\frac{1890}{3239}\frac{d_z}{d} > \delta \geq \frac{15}{28}\frac{d_z}{d}$	T_8
$\frac{3239}{1890}d > d_z > \frac{14\,037}{10\,750}d$	$\delta \geq \frac{1890}{3239}\frac{d_z}{d}$	T_6, T_7, T_8
$d_z > \frac{3528}{3239}d$	$\frac{1890}{3239} > \delta$	T_1