"Is the WTO's Article XXIV a Free Trade Barrier?"

Ben Zissimos and David Vines

CSGR Working Paper No. 49/00

February 2000

Centre for the Study of Globalisation and Regionalisation (CSGR), University of Warwick, Coventry CV4 7AL, United-Kingdom. URL: **http://www.csgr.org**

Is the WTO's Article XXIV a Free Trade Barrier?

Ben Zissimos and David Vines[1]
CSGR, Warwick, and IES, Oxford;
Balliol College, Oxford, Australian National University, and CEPR.

CSGR Working Paper No. 49/00
February 2000

Abstract:

Why is further multilateral trade liberalisation proving so difficult to achieve? This paper shows that Article XXIV itself, the set of WTO rules governing trade block formation, undermines the multilateral liberalisation process. Trade block formation under Article XXIV can be thought of as a coalition formation game with negative externalities. We suppose that the usual mechanism through which block formation exerts a negative externality on non-members - a rise in external tariffs - is precluded by Article XXIV. But essentially the same effect is created by internal tariff reduction. From this it follows that free trade is not an equilibrium.

Keywords: free trade, protection, trade liberalisation.

JEL Classification: F02, F13, F15.

Address for correspondence:

Ben Zissimos
CSGR
University of Warwick
Coventry CV4 7AL, UK

Email: benjamin.zissimos@economics.ox.ac.uk
Email: econdav@ermine.ox.ac.uk

[1] We thank Kyle Bagwell, Christopher Bliss, Francis Bloch, Paola Conconi, Ben Lockwood, Marcus Miller, Ray Riezman, Lucio Sarno, Bob Staiger, Margaret Stevens, Ted To, John Whalley, Myrna Wooders, Zhilian Yi and participants at the 1999 Midwest International Trade meeting, the 1999 MMF conference, University of Oxford, as well as seminars in CSGR (Warwick) and Nottingham University for useful comments and conversations about this paper. Any errors are our own. Financial support from the ESRC Centre for the Study of Globalisation and Regionalisation (CSGR), University of Warwick, is gratefully acknowledged.

Non-technical Summary

Why is further multilateral trade liberalisation proving so difficult to achieve? This paper investigates the idea that Article XXIV, the WTO's own set of rules on trade block formation, is at least partly to blame. The conventional view is that Article XXIV ensures that trade block formation facilitates free trade. Trade block formation can be separated into two parts. One is where members agree to remove all protectionist trade measures. The other concerns the setting of external tariff policy. Trade block formation can raise the power of its members on world markets; power that is exercised by raising external tariffs. Article XXIV stipulates that members, upon forming a trade block, are not allowed to become more protectionist towards non-members. So trade block formation must be trade liberalising overall. Providing that trade block formation leaves members better off, so the argument goes, this process can be expected to continue until all countries are a member of one all encompassing trade block, equating to free trade.

Whilst this argument is appealing in its simplicity, many believe that it is wrong, and that free trade will not be an equilibrium when countries can form a trade block, even under the conditions of Article XXIV. (Bhagwati 1993, Goto and Hamada 1998, and McMillan 1993 argue this, with reference specifically to Article XXIV; Bagwell and Staiger 1998, 1999 discuss the problems that can be created by preferential trade agreements more generally, citing those formed according to Article XXIV as a particular case.) However, past work does not allow the structure of trading arrangements, preferential or otherwise, to be determined as an outcome of the model, under the incentive structures created by the conditions of Article XXIV. It is the purpose of this paper to do exactly this, and show that the structure of trading arrangements that emerges is not characterised by free trade.

We show that the conventional view goes wrong by neglecting the fact that, even if blocks do not become more protectionist against non-members, trade block formation makes non-members worse off at the same time that it makes its members better off. This can be thought about most easily as an exchange rate effect (although in our paper, and in the wider trade theoretical literature, it is referred to as a terms of trade effect). Take, as an illustration, the efforts being made within the European Union (EU) to form a single market. The principle intention of this is to increase

trade between member states. With all else equal, each EU member will trade less with the rest of the world as a result. Consequently, their currencies will appreciate against those of non-Europeans like the yen and the US dollar. This increases the purchasing power by Europeans of non-European goods, making them better off. Conversely, it reduces the purchasing power of non-Europeans over European goods, to their detriment. All this happens because trade within the Union is facilitated, and happens even if trade with outsiders is made no more difficult.

How does this relate to whether trade block formation will lead to global free trade? The key insight is simple. If the benefits from membership of an exclusive club are derived partly by making outsiders worse off, then the club will not throw open its doors to all comers. Facilitating trade between block members has exactly this effect. The purchasing power of block currencies increases, whilst that of outsiders declines. Consequently, trade blocks do not have an incentive to allow all applicants to join, because some of the benefits of membership come from being able to purchase the products of outsiders more cheaply on world markets. So there is a limit to the expansion that can be expected from existing blocks, and free trade between all countries will not arise.

There will be those who say that some liberalisation is better than none at all. On the face of it even the partial liberalisation created by trade block formation under Article XXIV might appear to be a positive achievement. However, while liberalisation amongst members of a larger block enhances their purchasing power on world markets, outsiders loose purchasing power and become worse off. Therefore, trade block formation is a force for growing world inequality, even under Article XXIV.

In conclusion, there are many reasons why countries might have seemed more interested in becoming part of a trade block than engaging purely in multilateral trade liberalisation. The surprise is that WTO officials, in trying to understand why this is happening, should look to the incentives created by their own rules for at least part of the answer.

1. Introduction

One of the most important international economic issues at the moment is why the World Trade Organisation (WTO) is having such difficulty promoting further multilateral trade liberalisation. It is widely believed that the Uruguay round, the latest set of world trade negotiations completed in 1994, took too long and achieved too little given the resources expended. And general enthusiasm for another world trade round remains week. Yet many believe that the world would benefit from further liberalisation.

The purpose of this paper is to argue that Article XXIV, the WTO's own set of rules governing trade block formation, is undermining the multilateral trade liberalisation process. We do not pretend to be the first to propose this view. Most of the elements of our argument exist elsewhere in the academic literature. But this is the first occasion on which a model is set up where the block structure that emerges under the Article XXIV is determined endogenously, and shown not to be characterised by free trade.

The key insight that we make use of to prove our point is quite simple. It is that if the benefits from membership of an exclusive club are derived partly by making outsiders worse off, then the club will not admit as members all those who would like join. At its most basic level this may work on nothing other than the principles of kudos and envy. It becomes particularly interesting to economists when the mechanism through which insiders benefit and outsiders are made worse off operates through the market.

The connection has already been made by Yi (1996, 1997), among others, that a customs union can be thought of as a coalition of countries. The idea that customs union formation does harm to outsiders whilst benefiting insiders is generally accepted. The literature on coalition formation establishes, as a general principle, that if an exclusive coalition exerts a negative externality on non-members then it will not expand to embrace all individuals. The result that a customs union will not expand to free trade is immediate.

Models of customs union formation are regarded as controversial, because they assume that members jointly manipulate common external tariffs to maximise their own welfare. Suppose that trade block expansion mandates a rise in common external tariffs. Critics of this literature argue that external tariffs cannot rise, as might be predicted by these models. This is because such policy

action is precluded by Article XXIV, which stipulates that upon the formation of a trade block members must not raise the level of protectionism against non-members. And indeed, external tari¤ rates have not risen over the recent period of regionalisation[1].

If the mechanism through which the negative externality works is not allowed, in this case by Article XXIV, we are bound to ask whether the insights of the coalition formation literature still hold. Can we still argue that trade block formation disrupts the multilateral liberalisation process? To do so in a fully general way we must …nd some alternative means through which it exerts a negative externality on non-members.

We do indeed identify an alternative. Adapting a more general framework originated by Mundell (1964), we show that a terms of trade externality is created when block formation entails nothing other than the removal of internal barriers. The terms of trade impact on members is positive, while on non-members it is negative. Most appealing about this motivation is that it must, by de…nition, be a feature of all trade block formation. Although Mundell's original framework has no capacity for welfare analysis, we are able to analyse the welfare implications of this type of externality using a model of Dixit-Stiglitz (1977) relative demands. By being allowed to form an exclusive club under WTO rules, members improve their own welfare at the expense of others', via the improvement in block terms of trade. Trade blocks do not have an incentive to allow all applicants to join, because some of the bene…ts of membership come from being able to purchase the products of outsiders more cheaply on world markets. So there is a limit to the expansion that can be expected from the trade blocks that are formed, and free trade will not result from this process.

In principle, WTO rules exist in the form of Article XXIII that enable non-members to seek redress for the damage done to them when a trade block is set up (see Bagwell and Staiger 1999). Under this Article, governments can make a 'non-violation' complaint when their country's trade has been adversely a¤ected by a trade agreement reached by other governments, even if the agreement is consistent with the Articles of the GATT treaty. However, if this rule were fully e¤ective in practice, then we would have observed limited change in trade patterns as a result of the trade block

[1] The question of what conditions are required for trade block formation to mandate a rise in external tari¤s provokes an important set of issues in itself. These have been analysed comprehensively by Syropoulos (1999). As shall become clear, in this present paper we will focus exclusively on the situation where external block tari¤s would rise if this were not precluded by Article XXIV.

formation that has taken place over the last ten years. In fact, Soloaga and Winters (1999) provide empirical evidence to support the general perception that trade block formation has increased trade between members, at the expense of non-members, from which the terms of trade e¤ects that we identify will follow.

On the other hand, some argue that Article XXIV is being circumvented using methods of protectionism that are less visible than tari¤s, such as anti-dumping duties, e¤ecting a de facto rise in external barriers through trade block formation (see for example Panagariya and Gupta (1998). These e¤ects would clearly be complementary to those that we demonstrate. Our aim is to show that even if trade blocks have been formed in accordance with WTO rules, and external protectionism has not risen at all, trade block formation will not lead to free trade.

The theoretical literature on trade block formation under Article XXIV is surprisingly sparse[2]. One paper that has been written on this subject is by Syropoulos (1999). He sets up a model for looking at the impact on tari¤ setting, terms of trade and welfare of trade block formation. This general framework allows trade block formation under a variety of regimes to be analysed, enabling the requirements of Article XXIV to be examined as a particular case. He too shows that trade block formation under Article XXIV exerts a negative externality on non-members, whilst improving the welfare of members. However, because he wishes to examine aspects of a speci…c block structure under a variety of formation rules, he imposes the block structure exogenously in a 3 £ 3 model. In this present paper we show that the block structure forms endogenously as a sub-game symmetric equilibrium in an n £ n model. Generalising the number of countries introduces an important extra dimension to the analysis. The constraint in a 3 £ 3 model is that once two countries have formed a block, the best reply by the rest of the world in terms of block formation cannot be analysed, being ruled out by the fact that it is a single country.

The paper proceeds as follows. In the next section, we set out the general structure of the model, which draws on Krugman (1991) and Bond and Syropoulos (1996a). Within this structure, an analytical representation for the optimal tari¤ can be determined in a 'prisoners dilemma' type

[2] The implications of Article XXIV have been considered indirectly by Bagwell and Staiger (1998, 1999), as an instance where the two GATT/WTO pillars of reciprocity and non-discrimination are violated, so that in equilibrium a world welfare maximum cannot be guaranteed. Its implications are also discussed by Bhagwati (1993) and McMillan (1993). But none of these authors formalise trade block formation under Article XXIV, and none examine an equilibrium under its conditions.

world trade equilibrium, that would prevail in the absence of the WTO. The level of welfare under this solution can then be used as a benchmark against which the bene…ts of trade block formation can be compared.

The model is set up in such a way that countries would wish to raise tari¤s upon forming a block if they were allowed to do so. But we assume that Article XXIV prevents this, and that countries are forced to cap external tari¤s at the unilaterally optimal rate. Because the model structure and optimal tari¤ rate can be expressed analytically, we can explore the welfare implications of block formation in a formal way. This framework is used to demonstrate the terms of trade externality of trade block formation; members gain, non-members lose[3].

The analysis of Section 3 assumes that one group of countries forms a block while the other group continues to set policy unilaterally. In Section 4 it is assumed that all countries act on the incentive to form a trade block under Article XXIV. The trade block formation process is characterised as a coalition formation game. The focus of attention is on an equilibrium where all countries are a member of one of two blocks. This solution is analysed to show why the two blocks will not merge to form a single one, synonymous with global free trade. The overall level of world welfare is shown to be lower than the free trade level[4].

We also discusses the alternative of 'Open Regionalism'. Under these rules free trade is an equilibrium, and world welfare is maximised. But when the two regimes are allowed to coexist, as is the case in the world that we live in, we argue that the equilibrium under Article XXIV will prevail. Therefore, any proposal for reform of these rules must recognise that some countries are worse o¤ under 'Open Regionalism' than under Article XXIV. Conclusions are drawn in Section 5, setting the insights o¤ered by our research in the context of other recent research in this area.

[3] It is important to distinguish our argument from Viners (1950), that block formation can reduce e¢ciency, even for members, if it diverts the pattern of trade away from that dictated by comparative advantage. The analysis of Section 3 a¤ords us the opportunity of showing that this possibility is ruled out by the set-up of our model.

[4] A world welfare maximum is consistent with each government's objectives under free trade in our model because the sole aim of each government is to maximise national income. Grossman and Helpman (1994) show that this need not be the case if governments objective functions include other things. We do not examine such possibilities here. The more general point has already been made by Bagwell and Staiger (1998, 1999) that an e¢ciency loss is brought about by the presence of a terms of trade externality, whatever governments' wider objectives.

2. A Model of Trade with Unilaterally Optimal Tariffs

This section sets out the basic economic structure of the model. This structure is then used to determine the optimal tari¤ of countries acting unilaterally, in a 'prisoners dilemma' world trade equilibrium. When there are no rules to govern world trade, the best that any country can do is to assume that all trade partners will impose optimal tari¤s on all imports, and behave in the same way itself. In the absence of a multilateral trade organisation like the WTO, all attempts to form a trade block break down because of the usual incentive to cheat, that mean trade negotiations can be characterised as a 'prisoners dilemma'. Two purposes are served by this. One is to provide a benchmark, against which the merits of WTO membership can be compared. The other is to determine the external tari¤ rate that countries use if they form a trade block according to WTO rules.

2.1. The Economic Structure.

There are assumed to be N countries $i \in N \{1; ::; i; j; ::; n\}$, a …nite, non-empty set. Consumers and governments are denoted by their corresponding country identi…er (as $i \in N$). These countries are symmetrical with respect to their economic structure. But outcomes are not necessarily symmetrical. Any di¤erence between countries in the trade policy regime that they operate will be re‡ected by variation in their terms of trade. A model of pure exchange is used here, in order to keep things simple. Each country is assumed to be endowed with a particular variety of a horizontally di¤erentiated product. Country i is assumed to be endowed with a quantity x_i of good i. In order to abstract from the complexities of di¤ering country size we assume that all countries are endowed with the same quantity of output, and this is normalised at unity; $x_i = x_j = 1; 8i; j \in N$[5].

These assumptions may at …rst sight appear to imply a retrograde step, given recent research in this area. Bond and Syropoulos (1996a) specify a model with a general endowment structure, where all countries are endowed with some of every good, but that each country is uniquely endowed with an extra amount of one of the goods. This could be thought of as capturing the workings of a production model with the property that each country has a comparative advantage in the production of a single product. Our assumptions imply a special case of this more general form, in

[5] It has been shown that larger countries may loose out when forming a trade block with smaller ones (see Kennan and Riezman 1988). By abstracting from relative country size we are able to focus here entirely on the welfare e¤ects of the block formation process.

which general endowments are set to zero, and the endowment of the good in which the country has a comparative advantage is set to unity.

We want to focus on the speci…c case where trade block formation and expansion would bring about a rise in tari¤s, were this to be permitted. Bond and Syropoulos show that this will always be the case for the endowment structure of our model, which is common to that of Krugman (1991). Under this assumption, we can say unambiguously that Article XXIV will always impose a binding constraint on the external tari¤s set by trade blocks.

One way to motivate trade policy intervention is to suppose that consumers in all countries value the di¤erence between products over which countries have monopoly power on world markets. To introduce this property to the model, CES preferences are speci…ed, entailing demand for any good i which is less than perfectly elastic. The utility function is based on Dixit and Stiglitz (1977):

$$U_i = \left[\sum_{j \in N} c_{ij}^{\mu} \right]^{\frac{1}{\mu}}; \quad 0 < \mu < 1; \tag{2.1}$$

where U_i is the utility of the representative consumer in country i. This depends upon c_{ij}, the demand by a consumer in i from country j[6]. Assuming that the number of goods available for consumption is large, the elasticity of substitution implied by the C.E.S. utility function is $¾ = 1/(1 - \mu)$ and given the restrictions on μ this is …nite.

Given an ad valorem tari¤ imposed by country i on the imports of country j, denoted t_{ij}, consumers maximise utility (2.1) subject to their domestic budget constraint. From the solution to the consumer's problem, relative demands can be expressed in the following form;

$$\frac{c_{ij}}{c_{ii}} = \left(\frac{p}{(1 + t_{ij})} \right)^{\frac{1}{1-\mu}} \tag{2.2}$$

where c_{ii} is home demand for the domestic good, and $p = p_i/p_j$ is the domestic terms of trade. A rise in p brings about substitution towards imports, but a rise in a tari¤ reduces relative demand for the import that is a¤ected.

[6] By using the subscripts in this way, we establish a convention that will be used throughout the paper. Thus, the …rst subscript always denotes the country where a particular action is being undertaken; in this case consumption in country i. The second subscript refers to the country a¤ected by the action; in this case the origin of the goods being consumed.

Because imports are purchased on world markets, the national budget constraint is expressed in terms of world prices, and not those faced in the domestic market which may include a tari¤.

$$\sum_{j \in N} p_j c_{ij} = p_i \tag{2.3}$$

Given relative consumer demands for goods by (2.2), this constraint (2.3) determines the overall quantities that a country's endowments will buy on world markets. This emphasises that tari¤s are modelled as a transfer, which distort consumer decisions. The proceeds are redistributed lump sum to consumers, so that overall consumption levels are determined by the national endowment (normalised at 1) and terms of trade p_i.

2.2. The Dilemma of Trade Policy Outside the WTO.

. The optimal tari¤ rate can be derived quite conveniently, in the manner of Johnson (1953). Our method adopts the thinking behind Johnson's seminal work on this issue, and applies it to the model set up in this paper. Each government is assumed to set tari¤s so as to maximise welfare subject to the o¤er curve of each country with which it trades. This constraint ensures that world markets clear. It describes the o¤ers made by all other countries in exchange for country i's exports as terms of trade vary.

To derive the representative optimal tari¤ set by country i, we need an expression for the o¤er curve of the representative trade partner - country j. To simplify its speci…cation, we assume that all countries other than i set the same policy. This means that their goods will trade at the same price on world markets. So if we choose the good of country j to be the numeraire, we know that the good of some other country k will trade at the same price ($p_j = p_k = 1$). Therefore c_{jk} can be taken as representative of the imports that change hands between all other countries. However, c_{ji}; imports from country i; will di¤er to the extent that the tari¤ set by the government in country i is di¤erent to that adopted elsewhere.

The o¤ers that country j makes to country i depend on the preferences of the representative consumer. As a result of the assumptions just outlined, the utility of a consumer in country j take the following simpli…ed form:

$$U_j = \left[c_{jj}^{\mu} + (N - 2) c_{jk}^{\mu} + c_{ji}^{\mu} \right]^{1/\mu}. \tag{2.4}$$

The consumer problem can be solved in the standard way, using (2.4), to obtain an expression for import demand from i;of similar from to (2.2);

$$\frac{c_{jj}}{c_{ji}} = \left(p_i\left(1+\bar{t}\right)^{\phi}\right)^{\frac{1}{1-\mu}} \tag{2.5}$$

where $\bar{t}$ is the uniform tari¤ rate set by all other countries. The demand for imports from country i relative to imports from any other country, such as k, can similarly be expressed thus;

$$\frac{c_{jk}}{c_{ji}} = \frac{\left(p_i\left(1+\bar{t}\right)^{\phi}\right)^{\frac{1}{1-\mu}}}{\left(1+\bar{t}\right)^{\phi}} = p^{\frac{1}{1-\mu}} \tag{2.6}$$

Having determined relative demands, the budget constraint for country j is required to solve for the level of demand for each good. This is written as

$$c_{jj} + (N-2)\,c_{jk} + pc_{ji} = 1:$$

Using (2.5) and (2.6) in this constraint, and the international market clearing condition $c_{ij} = pc_{ji}$, the o¤er curve of country j with respect to imports from country i can be represented as;

$$c_{ji} = \left(\left(1+\bar{t}\right)^{\frac{\phi}{1-\mu}} + (N-2)\left(1-c_{ij}\right)^{\mu/(1-\mu)}c_{ij}\right)^{\frac{1}{\mu}} \tag{2.7}$$

This is just country j's import demand function with respect to country i. It can also be thought of as measuring the goods exchange that country j o¤ers to country i, given the tari¤ that it imposes on those imports $\bar{t}$. As the tari¤ rises, country j o¤ers country i a smaller quantity of exports for a given quantity of imports.

As Johnson explains, the welfare of country i is maximised when its welfare function is tangent to the o¤er curve of country j. At this point, domestic prices are equal to the gradient of the o¤er curve, found by di¤erentiating (2.7);

$$\frac{dc_{ji}}{dc_{ij}} = \frac{(1-\mu c_{ij})\left(1+\bar{t}\right)^{\frac{\phi}{1-\mu}} + (N-2)\left(1-c_{ij}\right)^{\mu/(1-\mu)}c_{ij}}{\mu\left(1-c_{ij}\right)c_{ij}} \tag{2.8}$$

The tari¤ rate that optimises welfare $t^{¤}_{ij}$ can then be calculated as the wedge between the domestic prices determined by (2.8) and the rate of exchange on international markets between these two goods; $dc_{ji}/dc_{ij} = \left(1+t^{¤}_{ij}\right)c_{ji}/c_{ij}$. Rearranging, the optimal tari¤ can be expressed as the

elasticity of country j's o¤er curve reduced by 1, which simpli…es quite conveniently as follows[7];

$$t^{¤}_{ij} = \frac{dc_{ji}}{dc_{ij}}\frac{c_{ij}}{c_{ji}} \; ¡ \; 1 = \frac{1 \; ¡ \; ¹}{¹(1 \; ¡ \; c_{ij})}: \tag{2.9}$$

The interesting thing to notice about (2.9) is that the optimal tari¤ set by country i does not depend directly on the tari¤ set by country j. This convenient analytical feature results from the underlying form of the CES preferences. It is also convenient that country i's optimal tari¤ is determined immediately in terms of its own import from country j. For the intuition behind this, we should think of c_{ij} not as imports by country i but as the o¤er of exports made by country j. A relatively high quantity o¤ered in equilibrium re‡ects a low demand elasticity ($¹ ! 0; ¾ ! 1$) for country i's good. Consequently, a tari¤ imposed by country i will have a relatively small negative impact on demand. This creates more scope to improve domestic terms of trade through the use of tari¤s, so the optimal rate will be relatively high. At the other extreme, as the consumer in country j …nds it increasingly easy to substitute away from country i's good ($¹ ! 1; ¾ ! 1$) the optimal tari¤ $t^{¤}_{ij}$ will tend to zero.

The speci…cation that we use here deliberately emphasises that the optimal tari¤ rate depends on the export of a single product. This is in marked contrast to the speci…cation used by Krugman (1991) and Bond and Syropoulos (1996a), where the optimal tari¤ depends upon the export share, potentially including a range of goods. In their work it was important to show that as a trade block expands the external tari¤ changes. In our present discussion, we wish to determine the (unilaterally) optimal rate for a single country that imposes the same tari¤ on each import from all its trade partners.

Symmetrical expressions can be derived for the o¤er curve of country i and optimal tari¤ setting for country i. These four equations can then be used to solve for the four unknowns, c_{ij}, c_{ji}, $t^{¤}_{ij}$, and $t^{¤}_{ji}$, to obtain a trade equilibrium relationship between countries i and j. This can be extended to obtain a world trade equilibrium.

As was the case for Krugman (1991) and Bond and Syropoulos (1996a), it is not possible in general to obtain an analytical solution for the optimal tari¤ $t^{¤}_{ij}$ and c_{ij} from these two equations.

[7] An alternative method can be used to determine the optimal tari¤, which involves calculating the export price elasticity, or equivalently the price elasticity of imports by the rest of the world. This is the approach taken by Krugman (1991) and Bond and Syropoulos (1996). We shall explain the intentions behind the approach that we adopt below, in contrast to that of earlier work.

Previous work in this area has resorted to numerical simulations in order to obtain solutions. In fact, it is possible to obtain an analytical solution, for certain speci…c values of μ. This we have done for two illustrative substitution elasticities ¾ = 2 and ¾ = 4. In both cases only one root is real, singling itself out for analysis. In each case, the root is extremely cumbersome, and displaying it would yield no economic insight. But the properties of each root can be illustrated using their corresponding numerical values.

These are displayed in Table 1. Having speci…ed the substitution elasticity, the only other parameter in the model is the number of countries N. We can see that N does not have to be very large for the value of the optimal tari¤ to approach the limiting case of $t^{¤}_{ij} = (1 ¡ μ)=μ$. When ¾ = 4, for example, $t^{¤}_{ij} ! 1=3$ (or 33%) as $c_{ij} ! 0.$. In fact, when N = 400, $t^{¤}_{ij}$ is within just over 1 percent of this value. And although imports are small, they are not de‡ected much from their free trade value; when $t^{¤}_{ij} = 0$, $c_{ij} = 1=400 = 0:0025$, whilst under this optimal tari¤ $c_{ij} = 0:00249$. This tells us that the tari¤ rate from the actual solution of equation (2.9) in equilibrium is approximated quite closely by the limiting result. Earlier we noted that the same method can be used to determine optimal tari¤s for all countries, so subscripts can be dropped, and optimal tari¤s set universally according to the expression

$$t^{¤} = \frac{1 ¡ μ}{μ} \tag{2.10}$$

We shall refer to the policy, where a country sets tari¤s on all imports at the optimal rate determined by (2.10) as Unilateral Optimal Protectionism (UOP). Of course, when all countries adopt the UOP regime, all goods will trade at the same price on world markets. This is because every country o¤ers the same quantity of exports in equilibrium.

Using (2.10) as an approximation has some very appealing advantages. For example, (2.10) can be used in the expression for relative demands (2.2) to measure the demand for imports relative to domestic goods in equilibrium;

$$\frac{c_{ij}}{c_{ii}} = μ^{\frac{¾}{1¡¾}} \tag{2.11}$$

Note that for $0 < \mu < 1$, $c_{ij}/c_{ii} < 1$, which confirms that the impact of the optimal tariff is to skew demand away from imports and towards the domestic good; under free trade $c_{ij}/c_{ii} = 1$. This makes the welfare analysis of the next section tractable.

Nothing substantive is lost from our discussion by assuming that there are a relatively large number of trading units; in our case countries. In the work of Krugman, (1991) and Bond and Syropoulos (1996a), the focus is to determine the optimal tariff when the number of trading units, in their case trade blocks, falls towards a single one that includes all countries. In our present context, because tariffs cannot rise when a block forms, we take advantage of the fact that the optimal tariff calculation has an analytical representation when the number of countries is large. In the dilemma equilibrium, then, all countries set tariffs on the imports of all imports from all other countries according to (2.10).

The optimal tariff representation that we derive is underpinned by the fact that there are no "small countries" in the sense of price takers in this model. All countries, no matter how small, impose an optimal tariff. They are motivated to do so by the monopoly power that is bestowed upon them through the endowment of a unique horizontally differentiated product, which consumers in other countries desire.

3. Trade Block Formation Under Article XXIV

We now suppose that the institutional arrangements are put in place, in the form of the WTO, that enable countries to form a trade block. Recent research by Maggi (1999) draws attention to the fact that the WTO facilitates international cooperation, rather than enforcing it. It has no power to punish those that violate the world trade rules that it presides over. But if the WTO provides independent verification that all signatories to a trade agreement are abiding by its conditions (and all are better off under the agreement than they otherwise would be) this can be sufficient to hold it together.

Developments in the coalition formation literature make it possible to analyse the equilibrium that will arise in a game where agents are far-sighted. One implication of far-sightedness is that players have no opportunity to cheat on their partners, as they are assumed to know immediately, and take retaliatory action. We suppose that the WTO provides an effective monitoring role, so all

members can verify that their trade partners are behaving in a far-sighted manner by sticking to the rules of the preferential trading arrangements that they join. Thus, there can be no short term bene…t to cheating, against which the long term cost of deviation must be weighed. Under these conditions an agreement will be self enforcing, given that it o¤ers a higher level of welfare than the alternative dilemma solution[8].

3.1. How the block is formed.

At its simplest, trade block formation according to Article XXIV entails a group of countries agreeing to remove internal barriers. This section de…nes such a group, and analyses what happens to the terms of trade between these countries and the rest of the world when they set up a trade block, assuming that non-members hold their policy choices constant.

Of the N countries in the world, N® form a block, while N (1 ¡ ®) countries continue to adopt a UOP regime. The value of ® is restricted to the range $1=N < ® < 1$ because there must be more than one, but at most all, countries in a block. The revised budget constraint for a country in the trade block is written as

$$p\left(c_{aa} + (N® ¡ 1)\, c_{ab}\right) + N\,(1 ¡ ®)\, c_{ay} = p: \tag{3.1}$$

We refer to two representative trade block members as countries a and b. Representative non-members, that adopt a UOP regime of continuing to set optimal tari¤s on all imports, are referred to as y and z. (We only ever have to refer to two countries of either type.) With regard to prices, we have already observed that with a single trade regime in operation across all countries, all goods are traded internationally at the same price. Now, with two trade regimes in operation - a trade block and UOP - two prices are needed. The international price of goods from countries outside the block is set as the numeraire ($p_y = p_z = 1$), and p is used to denote the terms of trade between block goods and non-block goods $p = p_a=p_y$.

[8] This assumption of complete information is required under the rules of the coalition formation game that we make use of. It has been shown that a minor weakening of this assumption can open the door to a much wider range of outcomes. Even if countries can cheat on their partners for only a short while before being observed, trade agreements may become unsustainable. Each government's rate of time preference becomes important in determining the sustainability of a trade agreement. The possibility of tacit coordination in trade agreements was introduced by Dixit (1987), and …rst examined in the context of trade block formation by Bond and Syropoulos (1996b). It has also been explored in terms of the viability of multilateral trade agreements when trade block integration deepens by Bond, Syropoulos and Winters (1999).

The revised budget constraint (3.1) shows that imports can arrive from inside or outside of the block. Demand for each will di¤er to the extent of variation in tari¤s. Using the appropriate Dixit-Stiglitz relative demands and the balance of payments constraint (3.1), the demand functions for the domestic good and imports of both origins can be derived. The demand function for the domestic good can be expressed as

$$c_{aa} = \left[1 + (N\theta - 1)\left(\frac{1}{1+t_{ab}}\right)^{\frac{\mu}{1-\mu}}\right]^{\frac{1}{1-\mu}} + N(1-\theta)p^{\frac{\mu}{1-\mu}}\mu^{\frac{1}{1-\mu}} \tag{3.2}$$

while for intra-block imports it takes the form

$$c_{ab} = \left[(1+t_{ab})^{\frac{1}{1-\mu}} + (N\theta - 1) + N(1-\theta)p^{\frac{-\mu}{1-\mu}}(\mu(1+t_{ab}))^{\frac{1}{1-\mu}}\right]^{\frac{1}{1-\mu}} \tag{3.3}$$

and for imports from outside the block

$$c_{ay} = \left[(p\mu)^{\frac{1}{\mu-1}} + (N\theta - 1)(p\mu(1+t_{ab}))^{\frac{1}{\mu-1}} + N(1-\theta)p^{-1}\right]^{\frac{1}{1-\mu}} \tag{3.4}$$

Comparison of (3.2) and (3.3) con…rms that when the block is formed, so that $t_{ab} = 0$, demand for the domestic good will be the same as for imports from other trade block members. On the other hand, from (3.3) and (3.4), demand for imports from inside and outside of the block will be the same when $1 + t_{ab} = 1 + t_{ay} = 1/\mu$ and $p = 1$, which replicates the situation in the dilemma solution. These demand functions can be used to calculate indirect utility for the representative trade block member.

The terms of trade between block members and non-block countries is determined so as to clear the world goods market. To work this out, demand conditions are also needed for non-block countries. The constraint on the demand of each non-member is determined by their national budget constraint;

$$c_{yy} + (N(1-\theta) - 1)c_{yz} + pN\theta c_{ya} = 1: \tag{3.5}$$

Notice that the endowment of the representative non-member y is equal to the numeraire, whereas the endowment of the block member a is priced in terms of p (see the right hand side of equation (3.1). This highlights the role of the terms of trade in a¤ecting the relative welfare of trade block

members to non-members. In general, trade block formation will have two effects on welfare. There will be static efficiency gains for members through tariff removal, and income effects through changes in the terms of trade.

The world goods market clearing condition is

$$N^2\theta(1-\theta)c_{ay} = N^2\theta(1-\theta)pc_{ya}. \tag{3.6}$$

Note that the terms capturing relative regime size (in N and θ) cancel so that $c_{ay} = pc_{ya}$; p equates the value of imports by block members from outside the block c_{ay} to the value of imports by non-members from inside the block c_{ya}.

Demand for imports from the block by non-members is derived from (3.5) and the relevant Dixit-Stiglitz relative demands as

$$c_{ya} = \left[(p/\mu)^{\frac{-1}{1-\mu}} + (N(1-\theta)-1)p^{\frac{-1}{1-\mu}} + pN\theta\right]^{-1}. \tag{3.7}$$

It will be useful later to determine the remaining elements of demand by non-block countries. The demand function for the domestic good,

$$c_{yy} = \left[1 + (N(1-\theta)-1)\mu^{\frac{-1}{1-\mu}} + N\theta p^{\frac{-\mu}{\mu-1}}\right]^{-1}; \tag{3.8}$$

and demand for imports by non-members from other non-members is represented by the function

$$c_{yz} = \left[\mu^{\frac{-1}{\mu-1}} + (N(1-\theta)-1) + N\theta p^{\frac{-\mu}{\mu-1}}\right]^{-1}. \tag{3.9}$$

In the dilemma solution, when all tariffs are set at optimal levels and terms of trade between all countries are equal, import demands by non-members from all other countries are the same; $c_{ya} = c_{yz}$. Because non-members continue to impose tariffs at the same rate on all imports, ($t_{ya} = t_{yz} = 1/\mu - 1$), the only route through which their demand for imports varies is through the difference in the terms of trade. As block goods become more expensive, consumers substitute away from them; as p rises above 1, $c_{ya} < c_{yz}$. This distinguishes non-members from block members, whose demands for goods from within and outside the block also vary to the extent of the difference in the internal and external tariff rates.

3.2. The Impact on the Terms of Trade of Lowering Internal Tari¤s.

We can now use our model to demonstrate that even if external tari¤s are prevented from rising, trade block members can shift terms of trade in their favour by removing internal tari¤s. The equilibrium relationship between the terms of trade p and the representative tari¤ rate between block members t_{ab} can be determined using (3.4) and (3.7) in (3.6);

(3.10)

$$t_{ab} = \frac{1}{p\mu}\left[\frac{1}{N\Phi - 1}\left\{\mu^{\frac{1}{\mu-1}} + N(1-\Phi) - 1\right\}p^{\frac{\mu}{1-\mu}} + N\Phi - (\mu p)^{\frac{1}{\mu-1}} - \frac{N(1-\Phi)}{p}\right]^{\frac{\mu}{\mu-1}} - 1$$

With all other tari¤s capped at the unilaterally optimal rate, the only variable through which the terms of trade can be altered is t_{ab}. By describing t_{ab} as the representative intra-block tari¤ we mean that all tari¤s within the block are assumed to be set at the same rate. This is important because the outcomes of co-ordinating policy in this way are di¤erent from those that obtain if t_{ab} were to be manipulated unilaterally.

To appreciate best the relationship between internal tari¤s and the terms of trade, look at Figure 1. The curve aa' illustrates the relationship de…ned by (3.10). The …rst thing to note about this relationship is that p = 1 when $t_{ab} = (1-\mu)/\mu$, the UOP rate (this can be veri…ed from (3.10). Next, observe that the curve aa' is downward sloping. Because all goods are assumed to be gross substitutes in this model, the curve aa' will be convex to the origin; this is a consequence of increasing opportunity cost of substitution between goods embodied in the CES utility function. Note also that the terms of trade gains are continuous in tari¤ reduction, increasing until they are removed completely. In fact, further terms of trade gains would be available if it were possible to subsidise intra-block trade; in other words as internal tari¤s move into the negative range. This is a standard second best result; the incentive to set tari¤s in one area of the economy is re‡ected in the incentive to subsidise in another area. For the purposes of this present analysis we assume, in keeping with WTO rules, that trade cannot be subsidised, and that tari¤s are the only instruments available to policy makers. Therefore, we say that the block is fully ‡edged when $t_{ab} = 0$.

The relationship illustrated in Figure 1 can be understood using a mode of analysis suggested by Mundell (1964). Suppose we start at the dilemma solution, where trade is balanced, and all countries adopt the UOP regime. Now what happens if country a agrees to form a block with

country b? We know that when country a lowers its tari¤ against the import from country b, if terms of trade are held constant the trade balance of b will move into surplus. If both countries reduce their tari¤s symmetrically against each other, we know by Cournot's law[9] that their combined surpluses at constant terms of trade must equal the total de…cit in the rest of the world. The terms of trade of countries a and b must rise relative to the rest of the world in order to restore balanced trade equilibrium. This e¤ect is general, and can be extended to encompass any number of countries. As all countries in the block lower tari¤s relative to one another, re‡ected by a fall in t_{ab}, their trade balances go into surplus with the rest of the world. A rise in block terms of trade p is required to restore equilibrium. Moreover, the terms of trade gained by members is increasing in the number of block-members. This is illustrated by bb'. If the curve aa' is drawn for any chosen proportion of block-member countries ® then bb' will correspond to some higher proportion $®^{0} > a$. The bb' curve shows that for a bigger block, there is a bigger rise in p for a given reduction of internal barriers. Why? The larger the number of members, the larger the trade surplus of the block with the rest of the world brought about by a given reduction in t_{ab}.

In addition, Figure 1 can be used to think about what would happen if two trade blocks were to merge. The process would be exactly the same as that for the merging of individual countries into a single trade block. As the two blocks remove mutual tari¤s this creates a trade surplus with the rest of the world, and a terms of trade improvement against non-members is required to restore balanced trade equilibrium.

Through the process of removing internal tari¤s, trade block members are able to improve their terms of trade on world markets in a way that does not violate WTO rules. A key corollary is that there is an externality associated with this process for non-members, as their purchasing power on world markets is reduced. As p rises, the purchasing power of non-member endowments, valued at unity, falls. The welfare implications of this shift in the terms of trade will be analysed in the next section.

The recent literature on regionalism focuses on externalities associated with external tari¤ co-ordination. Here we have drawn attention to the fact that there is a terms of trade externality associated with the removal of internal tari¤s as well.

[9] Cournot's law states that the sum of all trade balances across countries is identically equal to zero. This is di¤erent from Walras' law, which states that all excess demands within a single country must sum to zero.

3.3. The Welfare Implications of Trade Block Formation for Members.

In this sub-section, we show that the welfare effects for members of internal tariff removal are positive, operating via the terms of trade externality. The externality is shown to have a negative impact on non-members.

There is a key difference in the impact on welfare between members and non-members, of trade block formation. Members benefit from static efficiency gains when internal tariffs are removed, as well as a terms of trade effect. Non-member countries are affected only by change in the terms of trade.

Let us look first at members. To evaluate the welfare impact of tariff removal is difficult, because the utility function is not linear in prices and tariffs. If it were, we could evaluate the impact of tariff removal on welfare by solving the following differentiation problem;

$$\frac{dU_a}{dt_{ab}} = \frac{\partial U_a}{\partial t_{ab}}\bigg|_{\partial p=0} + \frac{\partial U_a}{\partial p}\frac{\partial p}{\partial t_{ab}}: \tag{3.11}$$

In the present set-up, it is only possible to solve (3.11) in the neighbourhood of the dilemma world trade solution, where we know that value of p (at p = 1) a priori. Having signed (3.11) at this point, the fact that this sign holds more generally must be established using simulation analysis.

If trade block formation is to be beneficial to members, the removal of internal tariffs must improve welfare; (3.11) must be negative. We already know from the previous sub-section that the final term, expressing the terms of trade effects of intra-block tariff change, is negative. It remains to show that the first term is also negative, and the second positive.

The first term measures the static efficiency gains brought about by tariff removal. Clearly, it is necessary to sign this term in order to fulfil the task at hand. But this exercise yields an additional insight, which is important here. It will confirm that the welfare effects that operate in this model do not arise due to the trade diversion effects that were first discovered by Viner (1950). His principle concern was that the lowering of mutual tariffs might actually reduce welfare; in the terms of this present discussion, $\partial U_a/\partial t_{ab} > 0$. This would happen if countries within the trade block were to switch their import demand away from the country with a comparative advantage towards a trade block member, whose production was less efficient, reducing overall efficiency, and

hence welfare. In signing this …rst term as negative, we prove that this possibility has been ruled out here.

The proof that $@U_a=@t_{ab} < 0$ is by demonstration. Using (3.2), (3.3) and (3.4) in (2.1), and di¤erentiating with respect to t_{ab} we obtain

$$\begin{aligned}\frac{@U_a}{@t_{ab}} &= \frac{1}{\mu}\{C_a\}^{\frac{1-\mu}{\mu}}\left\{\frac{\mu(N\alpha-1)}{1-\mu}\left(\frac{1}{1+t_{ab}}\right)^{\frac{\mu}{1-\mu}}c_{aa}^{\mu+1}\right.\\ &-\frac{\mu N(1-\alpha)(N\alpha-1)}{1-\mu}\left[p^{\frac{-\mu}{1-\mu}}\mu^{\frac{1}{1-\mu}}+(N(1-\alpha))^{-1}\right](1+t_{ab})^{\frac{-\mu}{1-\mu}}c_{ab}^{\mu+1}\\ &\left.+\frac{\mu N(1-\alpha)(N\alpha-1)}{1-\mu}(p\mu)^{\frac{1}{\mu-1}}\left(\frac{1}{1+t_{ab}}\right)^{\frac{\mu}{1-\mu}}c_{ay}^{\mu+1}\right\}\end{aligned} \tag{3.12}$$

where $C_a = \left[c_{aa}^{\mu}+(N\alpha-1)c_{ab}^{\mu}+N(1-\alpha)c_{ay}^{\mu}\right]$. The three lines on the right hand side capture the impact on welfare of changes in consumption of the home good and imports from within and outside the block that result from a change in t_{ab}. Notice that the …rst and last lines, capturing the welfare e¤ects of the change in domestic goods c_{aa} and imports from outside the block c_{ay} are positively signed. The second line, capturing the impact of imports from other block members c_{ab} is negative. If, as a result of lowering t_{ab}; the positive impact on welfare through an increase in c_{ab} dominates the negative impact from a fall in c_{aa} and c_{ay}, then (3.12) will be negative as we require.

The problem is made tractable by evaluating a small deviation from the dilemma world trade solution. At this point we know that $p = 1$, $1 + t_{ab} = 1=\mu$ and (from (3.3) and (3.4) $c_{ab} = c_{ay}$. Employing these restrictions, (3.12) simpli…es to

$$\frac{@U_a}{@t_{ab}} = \frac{1}{\mu}\{C_a\}^{\frac{1-\mu}{\mu}}\left\{\frac{(N\alpha-1)\mu^2}{1-\mu}\left[c_{aa}^{\mu+1}-\left(\frac{1}{\mu}\right)^{\frac{1}{1-\mu}}c_{ab}^{\mu+1}\right]\right\} \tag{3.13}$$

The sign of $@U_a=@t_{ab}$ is determined by the expression $\left[c_{aa}^{\mu+1}-(1=\mu)^{1=(1-\mu)}c_{ab}^{\mu+1}\right]$, as all other terms are positive. Although protectionism under the dilemma world trade solution implies $c_{aa} > c_{ab}$, it is straight forward to verify that $(1=\mu)^{1=(1-\mu)}c_{ab}^{\mu+1} > c_{aa}^{\mu+1}$, from which it follows that $@U_a=@t_{ab} < 0$. This establishes that a small reduction of internal tari¤s from the UOP rate with any number of other countries - from just one other trade partner to all others - will produce static e¢ciency gains. Our intention to rule out trade diversion as identi…ed by Viner is also veri…ed.

It remains to evaluate the sign of the second term on the right hand side of (3.11). In doing so, we seek to establish that the positive terms of trade e¤ects brought about by members' mutual removal of internal tari¤s has a positive impact on welfare. Furthermore we must show that these bene…ts cannot be internalised by individual countries when they set optimal tari¤s unilaterally.

To prove this proposition, the relevant point of departure must once again be the dilemma world trade solution, at which point all possible gains of unilateral policy setting have been exploited. The proof is very similar to the one that we have just seen. The welfare function (2.1) for the representative trade block member must be di¤erentiated, this time with respect to the terms of trade p, having substituted expressions for domestic and imported goods demands (3.2), (3.3) and (3.4). This gives

$$\begin{aligned}\frac{@U_a}{@p} = \frac{1}{\mu}\{C_a\}^{\frac{1-\mu}{\mu}} \Big[&-\frac{\mu^2}{1-\mu}N(1-\beta)\mu^{\frac{-1}{1-\mu}}p^{\frac{2\mu-1}{1-\mu}}c_{aa}^{\mu+1} \\ &-(N\beta-1)\Big[\frac{\mu^2}{1-\mu}N(1-\beta)p^{\frac{2\mu-1}{1-\mu}}c_{ab}^{\mu+1}\Big] \\ &+N(1-\beta)\Big[\frac{\mu}{1-\mu}\Big[\mu^{\frac{-1}{\mu-1}}+N\beta-1\Big]p^{\frac{2\mu-1}{1-\mu}}-\frac{(\mu-1)N(1-\beta)}{p^2}\Big]c_{ay}^{\mu+1}\Big]\Big]\end{aligned} \tag{3.14}$$

This time, note that the …rst and second lines are negatively signed, while the third line is positive. So an increase in the price of block goods a and b has a negative impact on the welfare of block members. In order for a rise in p to have a positive impact on welfare, we must show that the positive e¤ect of cheapening non-block goods such as those from y dominates.

Once again we make use of the fact that, at the dilemma world trade solution, $p = 1$, $1+t_{ab} = 1/\mu$, and $c_{ab} = c_{ay}$. Then (3.14) becomes

$$\frac{@U_a}{@p} = \frac{1}{\mu}\{C_a\}^{\frac{1-\mu}{\mu}}\Big[\frac{\mu}{1-\mu}N(1-\beta)\Big[(N-1)(1-\mu)+\Big[\frac{1}{\mu}\Big]^{\frac{-1}{1-\mu}}\Big]c_{ay}^{\mu+1}-\mu^{\frac{2-\mu}{1-\mu}}c_{aa}^{\mu+1}\Big]\Big] \tag{3.15}$$

Given, as observed earlier, that $(1/\mu)c_{ab}^{\mu+1} > c_{aa}^{\mu+1}$, that $c_{ab} = c_{ay}$ at the dilemma solution, that $(N-1)(1-\mu) > 0$, and $\mu^{(2-\mu)/(1-\mu)} < 1$ it follows that $@U_a/@p > 0$.

Comparison of (3.13) and (3.15) sheds light on the combined e¤ects through which trade block formation improves welfare. The static e¢ciency gains arise because consumption is more evenly distributed across goods. Increased welfare from increased consumption of imports from trade block

partners more than compensates for the fall in welfare from consuming less of the domestic good and imports from outside the block. The terms of trade e¤ect on welfare operates by increasing the purchasing power of block members over imports from outside the block c_{ay}.

Equations (3.13), (3.15) and (3.10) establish that for any group of countries that forms a trade block they can attain a rise in welfare through a small reduction of internal tari¤s ($dU_a=dt_{ab} < 0$). Unfortunately, we cannot conclude from the analysis that this holds for a large reduction, or complete removal of internal tari¤s ($t_{ab} = 0$), because of the non-linearities that are inherent in the model. But we can establish that this is the case by performing simulations. Figure 2 displays the combined e¤ect on welfare of internal tari¤ removal and its reciprocal impact on the terms of trade, for one particular set of parameter values; $N = 40$, $® = 0{:}3$ and $\mu = 0{:}5$.

This …gure illustrates that when a trade block forms, and mutual tari¤s t_{ab} are reduced from UOP rates to zero, the welfare of the representative member country a, U_a will increase throughout the process. The horizontal axis shows internal tari¤ rates. From (2.10) we know that when $\mu = 0{:}5$, the UOP tari¤ rate $t^{¤} = (1 ¡ \mu)=\mu = 1$ (1 = 100%). The vertical axis measures U_a. Figure 2 provides graphical con…rmation of our analytical result, found by signing 3.11), that $dU_a=dt_{ab} < 0$ in the neighbourhood of the dilemma world trade solution. But more than this, we can see that the welfare of members will continue to increase until internal tari¤s have been removed completely. The relationship is constructed using equilibrium $p ¡ t_{ab}$ combinations given by (3.10). Although Figure 2 is based on a speci…c set of parameter values, it is presented as an illustration of the general properties of the model. Its functional characteristics hold in general for all admissible parameter values. That is to say, it is downward sloping between $t_{ab} = 0$ and $t_{ab} = (1 ¡ \mu)=\mu$[10].

Also note from Figure 2 that welfare is not at a maximum even when $t_{ab} = 0$. Welfare would continue to rise if trade could be subsidised ($t_{ab} < 0$), but only up to a point, after which welfare would start to decline. This follows from the second best nature of the analysis, where block tari¤s are being removed while non-block tari¤s are left in place, which we discussed with reference to Figure 1.

[10]We have established this by plotting the relationship for a range of limiting parameter values, which can be obtained from the authors. Further details about the e¤ect of parameter variation on the welfare consequences of trade block formation will be discussed in the next section, when we analyse the world trade equilibrium.

Does welfare always increase with a rise in p when mutual block tari¤s are held constant ($@t_{ab} = 0$); in other words, does the positive sign of (3.15) hold for large shifts away from the dilemma solution? As we already know, the terms of trade can change for a country or members of a block even if they do not change their tari¤ policy at all. This happens if some other group of countries change their tari¤ setting policy.

The largest possible shift from the dilemma solution is to the point where mutual block tari¤s have been completely removed; $t_{ab} = 0$. Evaluating (3.14) at this point,

$$(3.16)\quad \frac{@U_a}{@p} = \frac{1}{\mu}\{C_a\}^{\frac{1-\mu}{\mu}}\left[\frac{\mu N(1-\Phi)}{(1-\mu)\,\mu^{\frac{1}{1-\mu}}+\Phi\left(1-\mu^{\frac{1}{1-\mu}}\right)}\right]\pounds$$
$$\left[(1-\mu)\mu^{\frac{1}{1-\mu}}+\Phi\left(1-\mu^{\frac{1}{1-\mu}}+\mu^{\frac{2-\mu}{1-\mu}}-\mu^2\right)\right]\left[N+N\Phi\left(\mu^{\frac{1}{1-\mu}}-1\right)-\mu\right]$$

Although (3.16) looks rather cumbersome, it is easy to check that it is positive for all admissible parameter values. The main reason for displaying it in this way is to show that p cancels completely from the right hand side. This is a very useful result. It tells us that an improvement in terms of trade will always improve welfare, for any given block size, when internal tari¤s have been completely removed. To illustrate the implications of (28) in a graphical way, we plot the relationship between the terms of trade p and the welfare of country a holding all tari¤s constant. This is illustrated in Figure 3, showing that welfare is generally increasing in the terms of trade; $@U_a/@p > 0$. There is intuitive appeal in this. We would expect a rise in income to improve welfare. And in this model the terms of trade determine income.

3.4. The Welfare Implications of Block Formation for Non-members.

What are the welfare implications of trade block formation for non-members? This is captured by di¤erentiating the expression for the welfare of non-members;

$$(3.17)\quad \frac{dU_y}{dt_{ab}} = \frac{@U_y}{@p}\frac{@p}{@t_{ab}};$$

which takes the same form as (3.11), but for the missing …rst term. This is because non-members leave all tari¤s in place, reaping no static e¢ciency rewards; $@U_y/@t_{ab} = 0$. The impact of trade block formation on the welfare of non-members comes about entirely as a result of a change in

their terms of trade. And given that we have just shown this e¤ect to be positive for members, it follows immediately that it is negative for non-members. The improvement in the terms of trade that lead to a rise in the welfare of members is re‡ected by a deterioration in the terms of trade for non-members. (The proof takes the same form as the one we have just seen in the form of equation (3.16), but because p appears in the reciprocal for non-members, the sign of each price change term is everywhere reversed). The relationship between the welfare of non-members U_y and the terms of trade p is given in Figure 4, for the same parameter values as used in Figures 2 and 3; N = 40, ® = 0:3 and µ = 0:5. We see that U_y declines as p increases. This is the case not just in the neighbourhood of the dilemma solution, but for all terms of trade p.

We know from Figure 1 that $@p=@t_{ab} < 0$, so the second term on the right hand side of (28) is negative. Taken with the …rst term the relationship is positive overall. A fall in mutual block tari¤s leads to a reduction in the overall purchasing power of non-members' endowments on world markets, leading to a fall in their welfare. The damage done to the welfare of outsiders will depend on the size of the block. The reduction in non-members' terms of trade when a small block forms will be slight. A larger block will have a larger impact, and the adverse consequences for welfare of being outside such preferential trading arrangements will become more severe.

The analysis of this section shows that a group of countries can improve their welfare relative to the dilemma solution, by forming a trade block according to Article XXIV. From this we can conclude that the dilemma world trade solution will not be an equilibrium.

4. World Trade Equilibrium Under Article XXIV

The discussion so far has focused on the incentives of one group of countries to form a block, holding the policy choices made by the countries in the rest of the world constant. This approach is su¢cient to show that the dilemma solution is not an equilibrium. But we have to go further in determining the world trade equilibrium itself, by assuming that all countries act on the incentive to form a block under the conditions of Article XXIV.

To do this, we draw upon recent developments in the literature on the noncooperative theory of coalition formation, particularly by Bloch (1996) and Yi (1996, 1997). The key insight that we use from these papers is that when the formation of an exclusive coalition exerts negative

externalities on non-members then the grand coalition, of which every player is a member, will not be an equilibrium.

For some time now it has been recognised that a trade block can be thought of as a coalition of countries. This idea was …rst formalised by Riezman (1985), using the core as the basis for equilibrium. The core, as an equilibrium concept, su¤ers from two main criticisms. One is that it requires a degree of cooperation amongst members which does not have clear noncooperative underpinnings. The other is that all agents, or countries in the present context, act simultaneously. The main dif…culty with this is that individuals cannot be farsighted, in that individual deviations cannot be countered by subsequent moves. These problems have lead to the formulation of sequential games of coalition formation where the process is described by an explicit extensive form non-cooperative game. As already mentioned, Yi (1996) has examined customs union formation using such a framework. He has shown how trade block formation gives members the power to improve their welfare through the manipulation of external tari¤s, to the detriment of non-members.

In the forgoing analysis of this present paper, we have seen how the removal of internal tari¤s can be used to the same e¤ect. And, as pointed out in the introduction, one of the most appealing aspects of this mechanism is that internal tari¤ removal must be a feature of all trade block formation.

The …rst task of this section is to set out the game of coalition formation. The model of the earlier sections is then adapted to describe the payo¤s to the coalition formation game. These payo¤s are used to show that the world trade equilibrium will support at least two blocks, implying that free trade will not arise.

Although the focus of this paper is on Article XXIV, in principle we can use our model to look at coalition formation under other rules of trade block formation. One such alternative is Open Regionalism. This set of rules requires that if countries form a block, it must be open to the membership of any country that would like to join. We examine the equilibrium that arises under these rules. We also discuss what happens when countries can choose between forming an Open Region or a block according to Article XXIV. This is novel because all previous (formalised) discussion on this matter has assumed that only one block formation regime would be on o¤er at a given time.

4.1. Article XXIV and The Coalition Unanimity Game.

In order use the insights of the literature on coalition formation, the key conditions of Article XXIV must be matched to the rules of a game that has been de…ned in the literature. The most appropriate type of sequential coalition formation game to characterise this set of rules is the "In…nite-Horizon Coalition Unanimity Game" of Bloch (1996). In this game, each player has the option to oppose a coalition that is proposed, and suggest an alternative. Thus each member of every coalition is maximising welfare; the equilibrium is consistent with each member's (non-cooperative) objectives.

Under the rules of this game, as under Article XXIV, members are allowed to form a block that is exclusive. They can agree to prevent other countries from joining. No country can force its way into a coalition. But at the same time no country can be forced to join. Bloch (1996) shows that providing a single player can obtain a non-zero payo¤ by acting unilaterally, and assuming that no payo¤ is received if no decision is reached on the coalition structure, there exists a subgame perfect equilibrium coalition structure of this game (see Bloch 1996, Corollary 2.5). These conditions are consistent with the model of this present paper. In particular, it will always be possible for a country to obtain a non-zero payo¤ because it can at a very minimum obtain a positive level of welfare by consuming its endowment under autarchy. The condition that no payo¤ is received until a decision is reached on coalition structure we now introduce by assumption.

Bloch (1996) goes on to show that when each player is ex ante identical (here the relevant parameter is player size) the game is symmetric. This is because a given coalition has the same value to its members, regardless of who they are. As a result, payo¤s depend only on the size of the coalition. The equilibrium under a game of this kind is denoted a symmetric subgame perfect equilibrium coalition structure (SSECS).

One of the main objectives of Bloch (1996) is to show that, under symmetry, the SSECS of the sequential game of coalition formation can be obtained as the equilibrium coalition structure of the much simpler "size announcement" game. The advantage of this type of game is that it highlights the structure of the coalition that is formed, playing down the importance of which players are the members. And, consistent with this, it is our intention to focus on the block formation process in assuming that all countries are ex ante identical in size. As a result, we are able to use the size announcement game to consider the process of trade block formation under Article XXIV.

To …t our model to the structure of this game, all countries must be ranked according to an exogenous order rule ½. We put country a …rst on the list, above country y: If each position is given a number, then a = 1, implying that a < y. This does not necessarily follow an alphabetical sequence; y does not necessarily equal 25, and z does not necessarily appear at position y + 1. We assume that the order rule could encompass any number of countries.

The game proceeds as follows. The …rst country on the list ½ - country a - announces the size of the coalition s_1 that it wants to form. Because all players are identical, the coalition size that is optimal for country a is also optimal for all its proposed block partners. So all proposed partners will accept a's o¤er. The members of this …rst block are taken from the s_1 ¡ 1 countries that follow a on the list ½. They then withdraw from the game. We assume that the next country on the list is country y. Country y then announces the size of its block s_2, and the formation of this second block follows the same process as the …rst. In general, this process continues until all countries are part of a block, or have opted to set policy unilaterally. Intuitively, this leads to a unique SSECS because each proposed coalition structure will be accepted by all partners immediately. Finally, note that no transfer of utility is allowed in the size announcement game. The robustness of the equilibrium coalition structures depends on this assumption. Therefore, the corresponding assumption is adopted here that there are no side-payments in the trade agreements that are reached. The implications of this assumption will be discussed below.

The aim is now to show how the terms of trade e¤ects discussed above will operate in conjunction with this coalition formation process to prevent the formation of a single block in equilibrium, thereby precluding free trade. The solution where there are just two blocks is characterised to verify that it is not individually rational for all countries to agree to the formation of a single block under the rules of Article XXIV. This two-block coalition structure is then checked for a range of parameter values to show that it can represent a world trade equilibrium.

4.2. A World Model of Two Blocks.

In this section, the payo¤s to members are examined when two blocks are formed according to the coalition formation process of the previous sub-section. The model of a representative block member a derived in Section 3 is used to consider the payo¤ of joining the …rst block. This consists of the budget constraint, (3.1) and the demand functions (3.2), (3.3) and (3.4). In discussing these

equations it was pointed out that when the block is fully ‡edged ($t_{ab} = 0$) the demand by a for its own good is the same as for imports from anywhere within its own block ($c_{aa} = c_{ab}$). So (3.2) and (3.3) are equivalent, and one of these can be dropped. The model of the representative non-member y is adapted to produce the model of the second block. This is done by lifting the assumption that y imposes a unilaterally optimal tari¤ on all imports. Instead, it now imposes tari¤s only on imports from the other block. The demand by country y for its own good or that of a block partner can be characterised by

$$c_{yy} = \left[N(1 - \alpha) + N\alpha\mu^{\frac{1}{1-\mu}} p^{\frac{\mu}{\mu-1}} \right]^{-1} \tag{4.1}$$

As with country a, demand for the domestic good is the same as for anywhere in the block ($c_{yy} = c_{yz}$). So all that is needed is an equation for imports from outside the block;

$$c_{ya} = \left[N(1 - \alpha)(p/\mu)^{\frac{1}{1-\mu}} + pN\alpha \right]^{-1} \tag{4.2}$$

From (3.4) (given that $t_{ab} = 0$) and (4.2), the terms of trade are determined entirely by the relative size of the two blocks. The relationship is given by

$$p^{\frac{1}{\mu-1}} + \frac{1-\alpha}{\alpha}\mu^{\frac{1}{1-\mu}} p^{-1} - \frac{1-\alpha}{\alpha} p^{\frac{\mu}{1-\mu}} - \mu^{\frac{1}{1-\mu}} = 0 \tag{4.3}$$

The term $(1 - \alpha)/\alpha$ measures the size of y's block relative to a's block. When there are just two blocks, both setting the same external tari¤s, relative size is the only variable in the determination of p, and the number of countries in the world N is immaterial to this relationship. Note that (4.3) does not have a general analytical solution. It is straight-forward to obtain an equilibrium $\alpha - p$ correspondence using numerical methods, for any speci…ed value of μ. However a set of roots can be found for $\mu = 0.5$; a substitution elasticity of $\sigma = 2$, as (4.3) takes the form of a cubic equation under this value. This enables us to explore the characteristics of the world trade equilibrium analytically. The root is non-linear, and rather cumbersome to write down, so we have put it in Appendix A. All we need to know from it is the equilibrium relationship that it describes between α and p. This is illustrated in Figure 5, as the line labelled p. Other solutions for $\mu \neq 0.5$ are also discussed based on the results of numerical simulations.

The horizontal axis of Figure 5 measures the share of countries in the block of country a, starting at ® = 0:5, or 50% of the world's countries, and increasing to 100%, constituting free trade[11]. There are two vertical axes. The one on the right shows the terms of trade p associated with a given share of country a's block. The one on the left registers welfare as a percentage of free trade; (free trade = 100).

Focus …rst on the relationship between trade block share and the terms of trade. When the blocks are of equal size, the terms of trade are equal at 1 (p is shown on the right hand axis). However, as the share of country a's block increases, terms of trade shift in its favour at an increasing rate. This illustrates the externality in the terms of trade e¤ect; for each marginal entrant, trade is liberalised with one more country than for the previous entrant, requiring a bigger adjustment in the blocks' terms of trade to restore equilibrium. This relationship is unde…ned when ® = 1, which explains why the line p stops short of the right hand side axis.

These equilibrium ® ¡ p combinations can be used to solve for the welfare of countries in each block. The U_a curve gives the welfare of country a, and all other members of a's block, using (3.3) and (3.4) in a CES utility function. The U_y curve shows the same for country y and all its block members. The welfare of membership is measured as a proportion of free trade on the left hand axis. Figure 5 presents the full range of outcomes in terms of welfare solutions, in much the same ways as Figure 4 of Bond and Syropoulos (1996a). However, here we are e¤ectively using them to describe the payo¤s of a coalition formation game. In the language of the coalition formation literature, Figure 5 shows the complete range of per-member partition functions. For all possible coalition structures ® in a two block world we have a mapping of what membership of each coalition is worth to each of its members.

When both blocks are the same size (® = 0:5) the welfare o¤ered by membership of either block is the same, at 90% of the free trade welfare level. However, as ® increases, the welfare o¤ered by country a's block rises above that o¤ered by the alternative. The welfare of a's block members will be maximised at ® = 0:9. The block share that maximises the welfare of each of a's members is denoted ®¤ (see Figure 5). For 0:5 · ® < 0:9 an increase in ® implies an improvement in welfare from the members of country a's block. Its terms of trade become more favourable, increasing

[11]We illustrate this range for ®, but the same set of arguments follow symmetrically for the same range of (1 ¡ ®).

the quantity of imports that members can buy from outside the block, thereby raising welfare. Meanwhile, y's members loose out through this process; as their terms of trade fall so do their level of welfare.

Notice that while country a's welfare achieves a maximum before all countries join the block, country y's welfare falls at an increasing rate as $\beta \to 1$. At β^{α}, the terms of trade gains with the rest of the world are exactly equal to the increased cost of importing from the marginal entrant. A member of a's block must purchase the goods of its fellow members not at the more favourable terms of trade with country y's block, but at parity. The members of a's block are better o¤ when they can purchase some imports more cheaply on world markets, than when they have to import all goods at the same price, as would be the case if they invited all countries to be a member of their block.

If the solution presented in Figure 5 were indeed an equilibrium, then in the …rst stage of the size announcement game country a would announce a block size of β^{α}, and country y's best reply would be to form a single block. The question of whether this two block solution is an equilibrium will be examined for a range of parameter values below.

Before hand, 'two block' solutions are characterised for two other substitution elasticities ϕ, and compared with one presented in Figure 5. These are presented in Table 2. (Results are presented to three decimal places where rounding would otherwise obscure the signi…cance of outcomes). When ϕ is relatively low, and consumers …nd it more di¢cult to substitute away from a given product, a given group of countries has more power on world markets. As a result, the ability to improve terms of trade through block formation is greater. For example, when $\phi = 1\frac{1}{3}$ the block size that maximises block welfare in country a is $\beta^{\alpha} = 0{:}92$, as against $\beta^{\alpha} = 0{:}9$ when $\phi = 2$ (the solution presented above) and $\beta^{\alpha} = 0{:}89$ when $\phi = 4$. This is essentially driven by the greater impact of the block formation process on the terms of trade when the substitution elasticity is relatively low. When $\phi = 1\frac{1}{3}$; $p = 2{:}69$ at a^{α}, which is signi…cantly higher than at β^{α} when $\phi = 2$. The value at β^{α} is lower still when $\phi = 4$. As was shown in Figure 5, the result of the terms of trade e¤ect is that at β^{α} welfare is higher for a member of a's block U_a than for a member of y's for each of the substitution elasticities ϕ: (It should be remembered that the level of welfare cannot be compared across substitution elasticities.)

These solutions characterise the outcomes for parameter values as they tend towards the extremes. As ¾ ! 1, the optimal block size does not rise signi…cantly above the value of ®¤ = 0:92. This is because the bene…ts of expansion due to the increase in the terms of trade are limited by the loss at the margin from having an increasingly small number of goods to purchase from outside the block at the relatively favourable terms of trade. On the other hand, as ¾ ! 1, ®¤ does not fall signi…cantly below ®¤ = 0:89. Although the terms of trade e¤ect of expansion become less powerful with an increase in ¾, which with all else equal reduces the bene…ts of block expansion, smaller blocks incur greater static e¢ciency losses because they impose tari¤s on a higher proportion of their total imports. These o¤setting e¤ects tend to keep the optimal size of the …rst block to form within fairly tight bounds around ®¤ = 0:9.

4.3. Demonstration of Equilibrium.

We now turn to the question of whether the two block solution represents an equilibrium. Yi (1997) derives conditions under which the equilibrium of a game with negative externalities is guaranteed to support just two coalitions, one larger than the other, as in the solutions presented above. However, it is also possible to show in a di¤erent way that, subject to the speci…cation of parameter values N and µ, the two block solution discussed here in Section 4 is an equilibrium. This is done by demonstrating that after the …rst block has been formed by country a at its optimal size ®¤, a subgame perfect equilibrium is brought about when country y proposes that all remaining countries form a single block. Each country maximises its welfare given the actions of all others.

Formally we check U_y when y announces $s_2 = N(1 ¡ ®¤)$ (given that $s_1 = ®¤$) against its payo¤s from announcing a smaller block. In fact, it turns out that we only need to check whether it is optimal to refuse entry to a single country. Intuitively, this is because the bene…ts to incumbents of accepting a marginal member are declining at the margin. As we shall see from the results of this exercise, if it is not optimal to refuse entry to a single country, it will not be optimal to refuse more than one.

For this to be possible, the model of two blocks must be adapted so that at least one country is not a member of either block, adopting a UOP regime. The representative single country is denoted ². The model has the same structure as set out above (see Appendix B for further details). The relative price of a single country's exports will be di¤erent to that of either of the blocks, and

this must be taken account of. A new relative price is introduced; $p_{a^2} = p_a/p_{^2}$. For consistency of comparison with the previous model of this section the other relative price $p = p_a/p_y$ is retained as before, with $p_y = 1$, and we continue to refer to the share of countries in a's block as ®. But now the share of y's block is expressed as $(1 - ® - {}^{-}/N)$; where $^{-}$ represents the number of single countries like ², setting UOP. We assume that the single countries adopt UOP because this gives members of y the highest possible level of welfare by excluding them. If the single countries were to form a block then terms of trade would shift in their favour and members of y, as outsiders would become worse o¤. So if it is not welfare maximising to exclude these countries when they adopt UOP, then it certainly will not be optimal to exclude them if they form a block.

There are now three market clearing conditions, of which we need two to solve for equilibrium. In terms of relative prices these can be presented as follows;

$$p = \left(c_{ay} + \frac{p}{p_{a^2}} c_{a^2}\right) = (c_{ya} + c_{^2a}) \tag{4.4}$$

$$p_{a^2} = (c_{a^2} + c_{y^2}) = (c_{^2a} + c_{^2y}/p)$$

For given demand functions (see Appendix B) the relative prices p and p_{a^2} can be solved using an iterative numerical search algorithm. Having determined the terms of trade, the level of welfare for a country under each of the regimes can be solved for.

Solutions to the model are presented in Table 3. The key result is that for each substitution elasticity ¾ the level of welfare for the representative member of y's block U_y is lower when a single country is excluded (Table 3) than when all countries are admitted (Table 2). For example, when ¾ = 2; and N = 40, U_y is 60:48% of the free trade level when one country is excluded, against 61.49% when all are admitted.

This result is driven by a terms of trade e¤ect, which can an be understood using Mundell's mode of reasoning once again. When country ² is ejected, it restores tari¤s against imports from y's block, and at constant terms of trade this will cause the trade balance in each of the countries in y's block to move into de…cit. A fall in block y's terms of trade with the rest of the world - equating to rise in p - is required in order to restore balanced trade equilibrium. This e¤ect can be seen by comparing p in Table 3, solved for each value of ¾; with the value in Table 2. Keeping to the example of ¾ = 2, and given that N = 40, p = 2:93 when ² is excluded, against 2:69 when

it is not. This means that goods from a's block, which make up the majority of every consumer's consumption bundle, are more expensive for members of y's block when ² is excluded, and being able to purchase less, they are worse o¤ as a result. From this set of results we see that for the parameter values presented it will be optimal for y to announce a block of size that includes all remaining countries when a has formed a block of size ®¤. Moreover, this choice of block size by y implies that country a and all its partners achieve a welfare maximum. Of course, country ² would be made even worse o¤ than country y if it were ejected from y's (compare $U_{²}$ with U_y in Table 3). This is because the de…cit that is created by ejection is concentrated within a single country, rather than being spread across a number of countries as in y's block. Consequently, the reduction in terms of trade is even larger. Therefore, it is also welfare maximising for a marginal country such as ² to accept y's o¤er of membership[12].

The results presented in Table 4 con…rm that if the rejection by y's block of a single country such as ² reduces welfare, then the rejection of two such countries reduces y's terms of trade with a's block, and hence welfare, further still[13]

From this simulation analysis we can conclude that, given the parameter values speci…ed for N and µ, the two block solutions analysed here represent a symmetric subgame equilibrium coalition structure.

4.4. Open Regionalism as an Alternative.

Open Regionalism has been adopted as an alternative set of rules for trade block formation by APEC. The key feature of Open Regionalism is that existing members of a trade block are not allowed to exclude others that would like to join. The meeting of APEC Heads of Government in Bogor in Indonesia in November 1994 gave regional approval to this policy as a regional strategy for block formation (see APEC 1994). As with exclusive block formation, this case has received careful analysis by Yi (1996, 1997), spanning both the game theoretical and international trade literature. The key result is that when no member can be excluded from a block that it wishes to

[12]The analysis by Syropoulos (1999) suggests that when a single country is excluded from preferential trading arrangements it may be welfare improving for it to reduce its external tari¤s. This happens because a country's terms of trade bene…t due to a given tari¤ declines at the margin as other trade blocks grow. However, as Syropoulos shows, this action will also improve the terms of trade and hence welfare of the block members, so can not result in ² obtaining a higher level of welfare than if it were invited by y to join.

[13]If this exercise is repeated for more countries, the losses in welfare for y's block relative to when all countries are admitted increase further.

join, every country will choose to join a single all-encompassing block in equilibrium. The coalition formation process under Open Regionlism is formalised by supposing that all players of the game announce an address, which corresponds to a coalition that they would like to join. For a game with negative externalities, it can be shown that all countries will announce the same address in the unique equilibrium.

To see this using our framework, using Figure 5, suppose that our model were used to describe the payo¤s of an Open Membership game. All countries are able to choose which block they would like to join. All countries would prefer to join country a's block than that of country y (this also being true for country y itself), because the payo¤ to joining the larger block is higher (given in Figure 5 by the gap between U_a and U_y). It is not possible, under the rules of the open membership game, to prevent a's block from expanding beyond the point that would yield optimal welfare of its members. So a single block will exist in equilibrium of which all countries will be a member, equating to free trade. This accords with the analysis of Yi (1997).

Thus outcomes vary according to the rules governing trade block formation. But what happens when a block can be formed according to either set of rules? In the real world, all countries o¤ered the choice of joining APEC also have the alternative of forming a block under Article XXIV. It is clear that a block of countries numbering 0:9N can do better by forming according to Article XXIV. So the equilibrium when a block can form according to either regime will be the same as that when only Article XXIV is o¤ered as a basis for trade block formation. This serves to highlight the problems that are likely to be associated with reforming Article XXIV. The analysis shows that it is not enough simply to o¤er Open Regionalism as an alternative. And as a majority of countries stand to do better out of trade block formation under Article XXIV than Open Regionalism, or equivalently free trade, it is di¢cult to see reform being achieved through majority voting.

A single block could be brought about, and overall welfare increased, if side payments between countries were possible. Then members of y's block could compensate members of a's block for the loss of welfare that would otherwise come about as a result of expansion to free trade. This raises the possibility of leaving members of y's block better o¤ and members of a's block no worse o¤. The possibility of welfare gains is brought about because all tari¤s are removed, entailing static e¢ciency gains. Therefore, the overall level of world welfare would be maximised, but of course the

original members of a's block would remain relatively well o¤, and those in y's block less well o¤ than if free trade had been achieved directly. While side payments of this kind are often allowed for in theoretical analyses of trade agreements, it is less clear that they are generally important in practice, although important exceptions are discussed by Whalley (1998).

5. Conclusions

In addressing the question of why the WTO is having such trouble in promoting the multilateral liberalisation process, we …nd that part of the answer lies in the incentive structure created by its own rules. The analytical framework of this paper shows that if external tari¤ levels are maintained at unilaterally optimal rates when a trade block is formed, welfare for members will increase, but at a cost to non-members. Through this e¤ect, Article XXIV prevents an equilibrium arising in which all countries belong to a single block, equating to free trade.

The incentive structure can be summarised as follows. Expansion in relative block size brings about a terms of trade gain for the members of the expanding block relative to non-members. But expansion reduces the number of goods (exported by non-members of the block) that can be purchased at the more favourable terms of trade. This e¤ect prevents the largest block that forms from expanding to include all countries.

For given parameter values, we show that equilibrium supports just two trade blocks, with all countries in the world belonging to one or the other. One block is relatively large, and its members enjoy a higher level of welfare than under free trade. The other is small, and its members are signi…cantly worse o¤. Because the members of the large block are better o¤ than under free trade, the majority of countries in the world have no interest in pursuing the multilateral liberalisation process. Under this unique equilibrium, world welfare is lower than the level that could be achieved under free trade.

The analysis of this paper questions the viability of Open Regionalism, when countries can choose between this as the basis for trade block formation and the rules of Article XXIV. We see that Open Regionalism o¤ers the free trade level of welfare. But block formation under Article XXIV o¤ers one group of countries a higher level of welfare still, and they are bound to choose this option if it is on o¤er.

What does this analysis tell us about the world that we currently observe? First, it suggests that the three regional blocks currently in existence are likely to become two, assuming that other costs and political barriers are not prohibitive. Secondly, APEC is likely to drop its commitments to the formation of an open region, and its members will opt instead to become part of exclusive trade block, keeping its external tari¤s in place.

The issue that cannot be addressed by this paper is the composition of these blocks; that of which countries will be in which block. Until the recent crisis in Asia, it seemed likely that APEC would merge with NAFTA. It was an open secret that many US policy makers would have liked to see this happen. But since the crisis, a certain amount of animosity towards the US appears to have developed in Asia, making such a merger unlikely, at least for the time being.

6. Appendix A

For the model of two trade blocks, the relationship between relative block share and terms of trade is expressed in the body of the paper as equation (4.3). This produces no generally valid expression for prices. The roots will vary according to the value of μ. Assuming a value of $\mu = 0{:}5$, three roots can be found using Mathcad, of which only one is real. Letting $\circ = (1 - \mathbb{R})/\mathbb{R}$, it can be represented as

$$\frac{1}{p} = \left(\frac{47}{96} - \frac{1}{1728}^{-3} + \frac{1}{288}\right)\left(\frac{1}{-16 + 6623^{-2} \cdot 16^{-4}} \cdot \frac{p}{3}\right)^{\left(\frac{1}{3}\right)} + \frac{\left(\frac{1}{12} + \frac{1}{144}^{-2}\right)}{\left(\frac{47}{96} - \frac{1}{1728}^{-3} + \frac{1}{288}\right)\left(\frac{1}{-16 + 6623^{-2} \cdot 16^{-4}} \cdot \frac{p}{3}\right)^{\left(\frac{1}{3}\right)}} - \frac{1}{12} \tag{6.1}$$

The root is plotted in $p - \mathbb{R}$ space in Figure 3.

7. Appendix B

Here the model Section 4.3 is set out, allowing for any number of single countries to be excluded when the second block is formed. The demand functions for country a, representative of its block

members, take the following form;

$$c_{ab} = \left[N\alpha + \frac{N}{p}\left(1-\alpha-\frac{1}{N}\right)(\mu p)^{\frac{1}{1-\mu}} + \frac{1}{p_{a^2}}(\mu p_{a^2})^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{ay} = \left[N\alpha(\mu p)^{\frac{1}{\mu-1}} + \frac{N}{p}\left(1-\alpha-\frac{1}{N}\right) + \frac{1}{p_{a^2}}\left(\frac{p_{a^2}}{p}\right)^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{a^2} = \left[N\alpha(\mu p_{a^2})^{\frac{1}{\mu-1}} + \frac{1}{p}N\left(1-\alpha-\frac{1}{N}\right)\left(\frac{p}{p_{a^2}}\right)^{\frac{1}{1-\mu}} + \frac{1}{p_{a^2}} \right]^{-1}.$$

The demand functions of country y are as follows;

$$c_{ya} = \left[pN\alpha + N\left(1-\alpha-\frac{1}{N}\right)\left(\frac{p}{\mu}\right)^{\frac{1}{1-\mu}} + \frac{p}{p_{a^2}}(p_{a^2})^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{yz} = \left[pN\alpha\left(\frac{\mu}{p}\right)^{\frac{1}{1-\mu}} + N\left(1-\alpha-\frac{1}{N}\right) + \frac{p}{p_{a^2}}\left(\frac{\mu p_{a^2}}{p}\right)^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{y^2} = \left[pN\alpha\left(\frac{1}{p_{a^2}}\right)^{\frac{1}{1-\mu}} + N\left(1-\alpha-\frac{1}{N}\right)\left(\frac{p}{\mu}\right)^{\frac{1}{1-\mu}} + \frac{p}{p_{a^2}} \right]^{-1}.$$

Finally, the demand functions of the representative singleton ² are:

$$c_{^2a} = \left[p_{a^2}N\alpha + \frac{p_{a^2}N}{p}\left(1-\alpha-\frac{1}{N}\right)(p)^{\frac{1}{1-\mu}} + \left(\frac{p_{a^2}}{\mu}\right)^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{^2y} = \left[p_{a^2}N\alpha(p)^{\frac{1}{\mu-1}} + \frac{p_{a^2}N}{p}\left(1-\alpha-\frac{1}{N}\right) + \left(\frac{p_{a^2}}{\mu p}\right)^{\frac{1}{1-\mu}} \right]^{-1};$$

$$c_{^2{}^2} = \left[p_{a^2}N\alpha\left(\frac{\mu}{p_{a^2}}\right)^{\frac{1}{1-\mu}} + \frac{p_{a^2}N}{p}\left(1-\alpha-\frac{1}{N}\right)\left(\frac{\mu p}{p_{a^2}}\right)^{\frac{1}{1-\mu}} + 1 \right]^{-1}.$$

The payo¤s to the representative members a and y of each block, and to the excluded country (or countries) ² are obtained by substituting the equilibrium demands into a CES utility function, as described in Section 4.

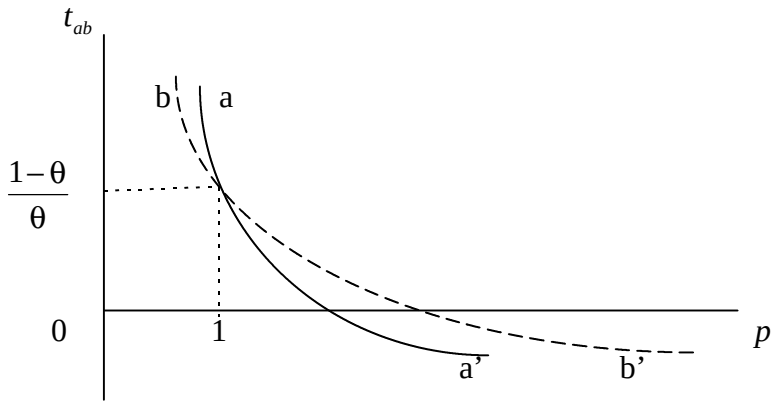


Figure 1

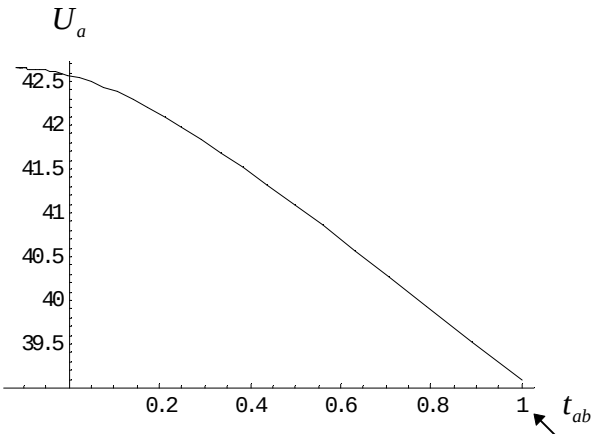


Figure 2

UOP: $t_{ab} = (1-.5)/.5 = 1$

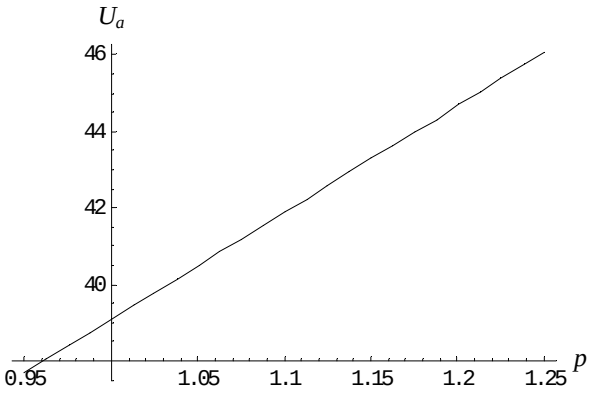


Figure 3

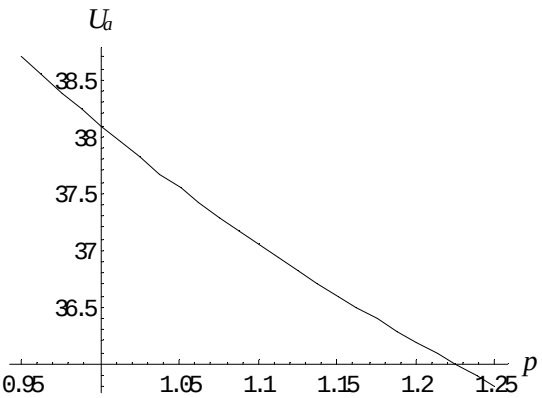
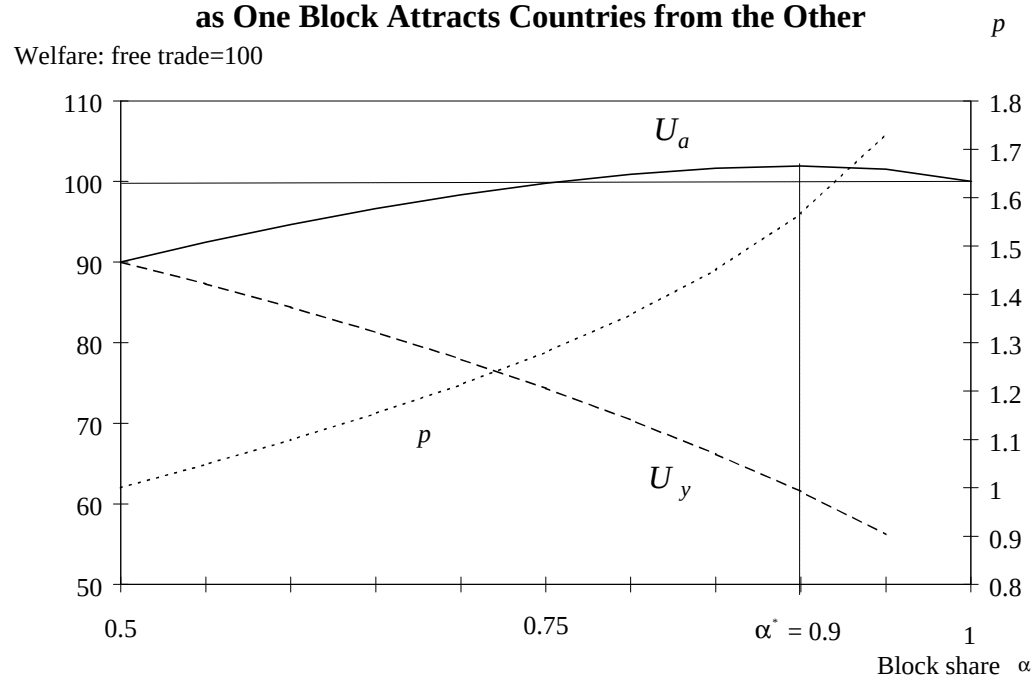


Figure 4

Figure 5: Terms of Trade and Welfare
as One Block Attracts Countries from the Other



World Welfare

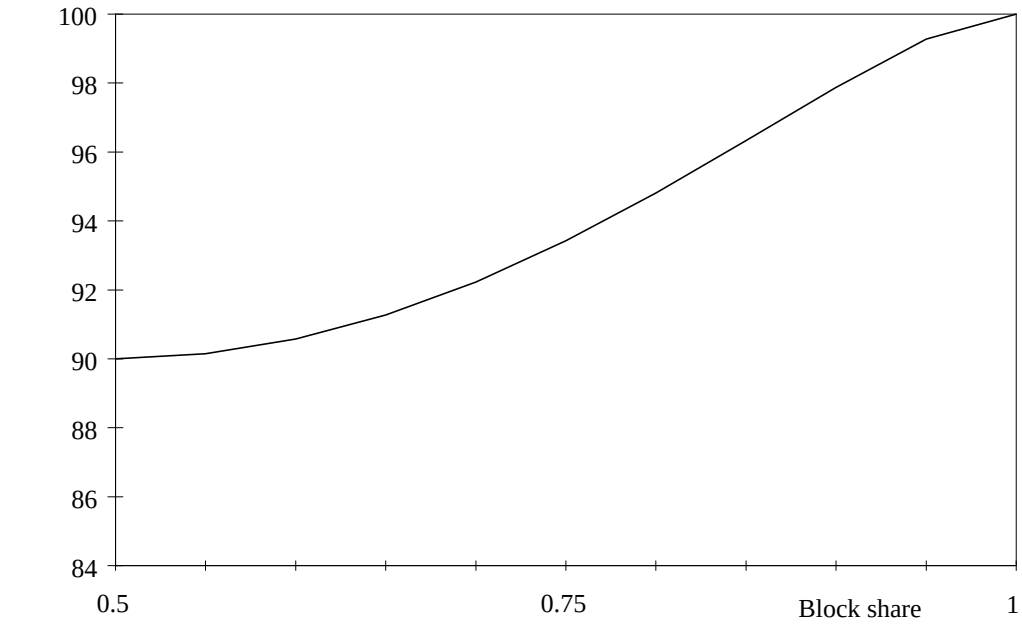


Table 1: Optimal Tariffs and Import Demand				
Substitution elasticity		$n = 4$	$n = 40$	$n = 400$
$\sigma = 2$	t^*_{ij}	1.151	1.024	1.00249
	c_{ij}	0.131	0.023	0.00248
$\sigma = 4$	t^*_{ij}	0.391	0.341	0.33416
	c_{ij}	0.148	0.024	0.00249

Table 2: Payoffs Under Two Blocks				
σ	α^*	p	U_a	U_y
$1\frac{1}{3}$	0.92	2.69	102.37	33.09
2	.9	1.569	101.93	61.49
4	0.89	1.192	101.00	83.15

Table 3: Payoffs Under Two Blocks when a Single Country is Excluded						
σ	α^*		p	U_a	U_y	U_ϵ
$1\frac{1}{3}$	0.92	$n = 40$	2.93	102.85	31.95	28.43
		$n = 400$	2.71	102.46	33.00	27.24
2	.9	$n = 40$	1.62	102.34	60.48	52.22
		$n = 400$	1.574	101.99	61.41	50.88
4	.89	$n = 40$	1.21	101.00	82.33	56.61
		$n = 400$	1.193	101.01	83.08	56.63

Table 4: Payoffs Under Two Blocks when Two Single Countries are Excluded						
σ	α^*		p	U_a	U_y	U_ϵ
$1\frac{1}{3}$	0.92	$n = 40$	3.26	102.37	30.45	29.39
		$n = 400$	2.74	102.54	32.90	27.39
2	.9	$n = 40$	1.69	102.31	59.07	53.23
		$n = 400$	1.58	102.05	61.33	51.05
4	.89	$n = 40$	1.23	100.84	81.28	56.02
		$n = 400$	1.20	101.02	83.01	56.66

8. References

APEC (1994) (Asia Paci…c Economic Co-operation) (1994) Achieving the APEC Vision: Free and Open Trade in the Paci…c, Second Report of the Eminent Persons' Group to APEC Ministers. Singapore: APEC Secretariat, Singapore.

Bagwell, K. and R. Staiger (1998) "Will Preferential Arrangements Underline the Multilateral Trading System?" Economic Journal 108(449) (July) pp 1162-1182.

Bagwell, K. and R. Staiger (1999) "An Economic Theory of the GATT," American Economic Review, 89(1) (March) pp 215-48.

Bagwell, K. and R. Staiger (1999) "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT," NBER Working Paper no. 7071, April.

Bhagwati, J. (1993) "Regionalism and Multilateralism: An Overview" J. de Melo and Panagariya (eds.) New Dimensions in Regional Integration, London: CEPR. Reprinted in Bhagwati, J., Krishna, P., and Panagariya, A. (eds.) 1999, Trading Blocks, MIT Press, Cambridge, MA.

Bloch, F. (1996) "Sequential Formation of Coalitions with Fixed Payo¤ Division and Externalities," Games and Economic Behaviour, 14: 90-123.

Bond, E.W. and Syropoulos, C. (1996a) "The Size of Trading Blocs, Market Power and World Welfare E¤ects", Journal of International Economics, 40, pp. 411-437.

Bond, E.W. and Syropoulos, C. (1996b) "Trading Blocs and the Sustainability of Inter-Regional Cooperation," in M. Canzoneri, W.J. Ethier and V. Grilli, (eds), The New Transatlantic Economy, London: Cambridge University Press.

Bond, E.W. and Syropoulos, C. and Winters L.A. (1999) "Deepening of Regional Integration and

Multilateral Trade Agreements," manuscript available at http://www....u.edu/~syropoul.

Dixit, A. (1987) "Strategic Aspects of Trade Policy", in Truman Bewley (ed.) Advances in Economic Theory: Fifth World Congress. Cambridge, Cambridge University Press, pp. 329-362.

Dixit, A. and Stiglitz, J.E. (1977) "Monopolistic Competition and Optimum Product Diversity," American Economic Review, 67: 297-308

Gros, D. (1987) "A Note on the Optimal Tari¤, Retaliation and the Welfare Loss from Tari¤ Wars in a Model with Intra-industry Trade", Journal of International Economics 23: 357-367.

Grossman, G.M. and Helpman, E. "Protection for Sale", American Economic Review 84(4), September, pp 833-850.

Johnson, H.G. (1953-4) "Optimum Tari¤s and Retaliation", Review of Economic Studies, 22: 142-153. (Reprinted as chapter 2 in H.G. Johnson, International Trade and Economic Growth, Allen and Unwin 1958.)

Kennan, J. and Riezman, R. (1988) "Do Big Countries Win Tari¤ Wars?" International Economic Review, 29(1): 81-85.

Krugman, P. (1991) "Is Bilateralism Bad?" in E. Helpman and J. Razin (eds.) International Trade and Trade Policy, MIT Press, Cambridge, MA.

Maggi, G. (1999) "The Role of Institutions in International Trade Cooperation," American Economic Review, 89(1): 190-214.

McMillan, J. (1993) "Does regional integration foster open trade? Economic theory and GATTs Article XXIV", in Anderson, K. and Blackhurst, R. (eds.) 1993, Regional Integration and the Global

Trading System, Harvester Wheatsheaf, Hemel Hempstead.

Mundell, R. A. (1964) "Tari¤ Preferences and the Terms of Trade," The Manchester School of Economics and Social Studies 32(1): 1-14. Reprinted in Bhagwati, J., Krishna, P., and Panagariya, A. (eds.) 1999, Trading Blocks, MIT Press, Cambridge, MA.

Panagariya, A. and Gupta, P. (1998) "Anti-dumping Duty versus Price Negotiations", The World Economy 21(8): 1003-1020.

Riezman, R. (1985) "Customs Unions and the Core," Journal of International Economics 19: 355-365.

Soloaga, I and Winters A. (1999) "Regionalism in the Nineties: What E¤ect on Trade?" CEPR Discussion Paper no. 2183, June.

Syropoulos, C. (1999) "Customs Unions and Comparative Advantage," Oxford Economic Papers 51 239-266.

Venables, T. (1982) "Optimal Tari¤s for Trade in Monopolistically Competitive Commodities," Journal of International Economics 11: 225-241.

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

Whalley, J. "Why Do Countries Seek Regional Trade Agreements?" in Frankel, J. (ed.) The Regionalisation of the World Economy. Chicago, University of Chicago Press, pp. 63-87.

Yi, S.-S. (1996) "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good" Journal of International Economics, 41: 151-175.

Yi, S.-S. (1997) "Stable Coalition Structures with Externalities," Games and Economic Behaviour, 20:201-237.