Increasing returns, multinationals and geography of preferential trade agreements

Rodney D. Ludema*

Department of Economics and School of Foreign Service, Georgetown University, Washington, DC 20057-1045, USA

September 2000

Abstract

This paper uses a model of horizontal multinational enterprises to explore the relationship between transportation costs and trade policy cooperation. Tariffs have the effect of attracting foreign direct investment to the benefit of consumers in the host country. As transport costs fall, the incentive to impose tariffs falls and the benefits to cooperation rise. Thus, in a repeated game in which cooperation is limited by a self-enforcement constraint, a reduction in transport costs facilitates free trade. This logic is applied to a three-country model to examine preferential trade agreements. It is found that if any country is too distant from the others, then global free trade is not attainable. Rather, if two of the countries are within a critical distance of each other and distant from the third country, then the unique outcome is an exclusive free trade agreement between the two adjacent countries. Thus, the model predicts a strong regional bias in preferential trade agreements.

Keywords: regionalism, trade agreements, multinational enterprises, economic geography.

JEL classification: F13, F15, F23, R12

*Tel: +1-202-687-1429; Fax +1-202-687-6102
Email address: ludemar@georgetown.edu (R. Ludema)

Acknowledgements

Thanks to James Anderson, Kyle Bagwell, Wilfred Ethier, Roger Lagunoff, Anthony Venables, Ian Wooton, and participants in several seminars for comments and discussions. I especially thank Robert Staiger and an anonymous referee for invaluable suggestions. Financial support was provided by a Georgetown University Junior Faculty Research Fellowship. Any errors are those of the author.

1. Introduction

Of all the variables that might explain (in a statistical sense) why countries enter into preferential trade agreements (PTAs), one stands out above all others: proximity. Virtually all PTAs are between geographically contiguous countries. Those that are not tend to be based on former imperial relationships, which have been diminishing in importance.[1] Indeed, so strong is the relationship between proximity and PTAs that few economists even bother to make a distinction between preferential trade and "regionalism."[2] Despite this obvious link, however, the theory of PTAs has not produced a compelling argument for why it exists. This paper suggests an approach to this issue and constructs a simple example, based on the theory of multinational enterprises (MNEs), to illustrate its potential.[3]

The approach taken in this paper will be summarized briefly at this point, postponing the details until the next subsections. There is a certain class of international trade models, involving elements of increasing returns and imperfect competition, in which transportation costs (which are closely correlated with distance) provide an argument for trade policy. Perhaps the most transparent model in this class is that of a MNE that faces a trade-off between proximity and concentration.[4] By restricting trade, an importing country can induce the MNE to substitute local production for exports, and this may benefit the country's consumers through lower prices. The desirability of a trade restriction for the importing country, therefore, may increase with the transport cost.

---

[1]An exception is the US-Israel FTA. Also, Norway and Switzerland are members of the European Free Trade Area (EFTA) but are not contiguous. However, all the countries that lie in between are members of the EU, which has a free trade agreement with the EFTA countries.

[2] See Baldwin and Venables, 1995, for a discussion of this point.

[3] Ethier (1999) has also addressed this issue, offering several alternative theories that emphasize the role of foreign direct investment.

[4] This terminology is borrowed from Brainard, 1993, 1997.

A branch of recent literature has sought to explain the nature of trade agreements by appealing to limitations placed on international cooperation by the absence of outside enforcement (e.g., Bagwell and Staiger, 1990; see Staiger, 1995, for a literature review). Without a supra-national enforcer, countries must design their trade agreements to be "self-enforcing." In a repeated game framework, this requires that at every point in time a country's one-time incentive to deviate from an agreement be less than the discounted benefit of future cooperation. Using the simple MNE model described above, this paper shows that the one-time incentive to deviate is lower and the benefit to cooperation greater the lower is the transport cost. Thus, countries separated by low transport costs are more likely to be capable of maintaining a free trade agreement than countries separated by high transport costs.

1.1. An argument from standard preferential trade theory

One of the earliest lessons from the standard theory of PTAs (Viner, 1950) is that a PTA is more likely to increase the welfare of its members if it includes the lowest cost suppliers of the goods being traded, as this eliminates the possibility of trade diversion. The cost of transporting goods across distance (as distinct from government-imposed tariffs) is one component of the cost of delivering a product to a foreign market. Other things equal, contiguous countries will be able to deliver goods to each other at lower cost than countries distant from one another. By the same token, a PTA between contiguous countries should be more likely to raise the welfare of its members than a PTA between distant countries.

Krugman (1991a) offers an extreme example. Imagine that transportation costs are zero between countries on the same continent and prohibitive to trade between countries on different continents, even without tariffs. It follows that an inter-continental PTA would offer no benefits at all, while an intra-continental PTA would liberalize all trade that can be liberalized and thus certainly raise the welfare of the continent as a whole. The continent in this example is said to be a "natural" trading bloc.

Although there is some question as to whether the notion of a natural trading bloc generalizes beyond simple examples (see, Frankel, et. al., 1997), the basic idea appears sensible. However, to make use of it as an explanation for the observed regional bias in PTAs,[5] one must make the additional assumption that governments actually enter into those trade agreements that give the highest welfare. While this may not be an objectionable assumption *per se*, it is a treacherous one in this context. If the policy environment in which trade agreements are created were so benign as to enable countries to achieve maximal welfare, then why wouldn't countries simply agree to global free trade and not bother with PTAs? A credible theory must account for the imperfections in domestic or international institutions that give rise to PTAs in the first place.

1.2. A counter-argument based on repeated game theory

While there are several approaches to modeling the institutional imperfections that give rise to trade policy, the most successful approach in analyzing the character of trade agreements has been that based on the theory of repeated games. In such a model, the limits of cooperation are determined by the balance of the one-time incentive for a country to deviate from an agreement with the discounted benefit of future cooperation (avoiding trade war).

---

[5] This is not to suggest that Krugman, or others who have put forth the natural trading bloc idea, have offered it as an explanation for the observed regional bias in PTAs. To the contrary, their focus has been almost entirely normative (i.e., on evaluating the regional bias in PTAs, not explaining it), as was Viner's.

Virtually the only argument in standard trade theory for *not* unilaterally eliminating all tariffs is the one based on monopoly power in trade.[6] The relationship between transport costs and monopoly power is generally ambiguous, but in most cases we would expect it to be negative. In Krugman’s example, there is no incentive for a country of any size to unilaterally impose a tariff on inter-continental trade, whereas a large country can expect to benefit from a small tariff on intra-continental trade through improved terms of trade. In other words, while there are greater benefits to intra-continental cooperation, there is a greater incentive to deviate as well. Thus, it certainly cannot be supposed that contiguous countries will be able to agree upon lower tariffs than distant ones, as long as we remain in a conventional constant returns-perfect competition framework.[7]

1.3. *An argument based on economies of scale*

The literature on trade policy under increasing returns and imperfect competition reveals many new channels by which trade policy can affect welfare, several of which depend on transport costs. Two papers by Venables (1985 and 1987) illustrate important cases. Venables (1985) examines a two-country model of oligopoly with free entry, homogeneous products and segmented markets, while Venables (1987) uses the Dixit-Stiglitz monopolistic competition framework. In each model, because of a transport cost, each firm sells more at home than it exports. A tariff imposed by one country causes firms to enter that country and exit the other. This has the effect of expanding domestic production to the benefit of domestic

[6]With the addition of political economy considerations, a country may stop short of complete unilateral liberalization, but the important point is that monopoly power in trade is the basis for trade agreements in a standard model. See Bagwell and Staiger (1999) for more on this point.

[7] There have been several papers to consider preferential trade agreements in perfectly competitive settings, using repeated games to model the limits of cooperation, e.g., Bond (1999), Bond and Syropoulos (1996) and Bagwell and Staiger (1997a and 1997b); however, only Bond (1999) includes transport costs. Bond demonstrates the possibility that a PTA between regional partners may be preferred to either a multilateral agreement or a PTA

consumers. Helpman and Krugman (1989) refer to this as the home-market effect. This effect appears in numerous other models as well, such as the geography model of Krugman (1991b).

The same effect can arise in models of MNEs; particularly those involving of horizontal multiplant production.[8] In such models, the decision to engage in multinational production reflects a trade-off between the firm's desire to be close to foreign markets (because of trade costs) and the desire to concentrate production at home and exploit economies of scale. Such models include Horstmann and Markusen (1992), Brainard (1993), Markusen and Venables (1998) and others (see Markusen, 1995, for a summary). Trade restrictions can induce firms to engage in multinational production ("tariff-jumping" foreign direct investment) which may benefit consumers in the host country as such production avoids transport costs (Smith, 1987; Motta, 1992).[9] All of this points to a general observation: the incentive to impose trade restrictions may be increasing in the transport cost in models where the home market effect is dominant.

In this paper, this positive relationship is exploited to generate a regional bias in trade agreements. Two cases are studied. The first is a two-country model that serves to illustrate the basic connection between transport costs and trade policy cooperation. It is shown that the benefits of cooperation rise and the incentive to deviate falls as the transport cost is reduced. There is also a discontinuity in the relationship between trade agreements and the transport cost. There is a transport cost threshold that must be crossed before cooperative tariff

---

between distant partners but finds PTAs between distant partners to have lower internal tariffs than those between regional partners.

[8] A horizontal MNE is a firm that sets up plants abroad to produce the same product that it produces at home. This is distinguished from a vertical MNE, in which divides different stages of the production process between plants in different countries. See Caves (1996) for details.

[9] Tariff-jumping FDI can also occur in a constant or decreasing returns to scale context (as in the original treatment by Horst, 1971), and, in such a context, it may be decrease welfare (Brecher and Daiz-Alejandro, 1977) or increase it (Srinivasan, 1983). It is only through scale economies and transport costs, that tariff-jumping FDI produces welfare gains through the home market effect.

reductions can occur at all, but once the transport cost falls below this threshold, free trade becomes supportable.

The second case is a three-country "triangular" model, in which two countries are adjacent to each other and equidistant from the remote third country. In this case, there is a transport cost threshold in each dimension. Below both of them, global free trade is attainable, as are most forms of preferential agreement. Above both of them, no cooperation is possible. If transport costs are low between the adjacent countries and high elsewhere, then the unique outcome is a preferential trade agreement between the adjacent countries. As is typical of repeated games, the less countries discount the future the higher will be the transport cost thresholds, meaning the easier it will be to sustain cooperation.

Taken together our results suggest that free trade will tend to arise between countries that are geographically near one another, and that preferential trade agreements, in particular, will be associated with regions (areas within which transport costs are low and between which transport costs are high).

In what follows, we demonstrate these points using a extremely simple model of horizontal MNEs, based roughly on Krugman (1983). While its primary virtue is tractability, this model accords with a number of well-known stylized facts (a feature it shares with most of the leading, though more complicated, models in its class). For example, foreign direct investment consists primarily of two-way flows between similar (developed) economies, in similar industries, and most of it is horizontal (Markusen, 1995). The decision of a firm to sell through a foreign affiliate rather than export is substantially affected by tariffs, transport costs, and scale economies (Brainard, 1997).[10] Moreover, there is some evidence that governments

[10] The evidence of tariff-jumping FDI is quite extensive (Caves, 1996). Recent evidence shows that even the threat of protection may induce FDI (Blonigen and Feenstra, 1997).

actually pursue trade policies for the purpose of influencing FDI. As Caves notes, "a great deal of survey and anecdotal evidence confirms the influence of tariffs on MNEs' locational decisions, not least because many trade restrictions have sought just that result"(Caves, 1996, p. 34). Australia and Canada are commonly cited cases (see, e.g., Krugman, 1992).

The rest of the paper is organized as follows. Section 2 sets out the basic MNE model. Section 3 examines the two-country case, introduces the repeated game framework, and demonstrates the relationship between transport costs and cooperation. Sections 3.2 and 3.3 discuss the robustness of the basic model. Section 4 examines the three-country model and establishes the main results regarding preferential trade agreements. Section 4.5 considers the effect of an alternative punishment mechanism. Section 5 concludes.

2. The Model

We consider a world consisting of $m \geq 2$ countries and two sectors, X and Y. There are n firms in each country, each producing a unique, firm-specific variety of good X from at least one plant. Each firm operates a plant in its home country from which it supplies its home market. We take the fixed cost of operating this plant to be zero. The firm may also operate plants in foreign countries. Doing so enables the firm to avoid the trade costs associated with exporting to those markets; however, each new plant incurs a fixed cost of F. The marginal cost of production is constant at c for all plants.

Trade in good X is subject to an iceberg-type transport cost, which causes goods to be lost or destroyed as they are shipped from one location to another. Between any two locations i and j, it is assumed that for every unit of good X shipped only $r_{ij} \leq 1$ units actually arrive. We assume that r_{ij} is equal to one between a plant and its local market and is strictly less than one between different countries.

Trade in *X* is also subject to *ad valorem* tariffs imposed by the governments. Tariffs are taken to be nonnegative and to apply on the basis of origin (where the good is produced) not ownership (who owns the plant that produces it). This ensures that when a firm locates a plant in a foreign country, the plant faces the same costs as all other plants located in the same country. Tariffs may discriminate on the basis of origin, as in the case of a PTA. PTAs are not required to maintain a common external tariff, so the analysis accommodates both customs unions and free trade areas. The structure of the game will be discussed in more detail later, but for now it is assumed that firms take tariffs as given when making their location and production decisions.

The representative consumer has a quasi-linear, separable utility function,

$$U = \sum_{i=1}^{mn} \theta C_i^{(1/\theta)} + Y, \qquad \theta > 1 \tag{1}$$

where C_i denotes consumption of variety *i* of good *X*, and *Y* represents consumption of the homogeneous, numeraire good *Y*. Each country is assumed to be endowed with an ample quantity of *Y*, and trade in this good is free. Differentiating (1) with respect to C_i gives demand functions $C_i = P_i^{-\varepsilon}$, where P_i denotes consumer price and $\varepsilon \equiv \theta/(\theta - 1)$ is the elasticity of demand.

2.1 Trade costs and market clearing

If a firm wishes to supply a market that is subject to *r*, it must produce a quantity 1/*r* times the quantity demanded in the market, so as to compensate for the losses in transit. Also, the consumer price in the market must be (1 + *t*)/*r* times the producer price, reflecting both *r* and any tariff rate *t* imposed by the importing government. Thus, to clear a market subject to *r* and *t*, a firm must produce a quantity

$$X(s,t) = (1+t)^{-\varepsilon} P_0^{-\varepsilon} (1-s) \tag{2}$$

where $s \equiv 1 - r^{(\varepsilon - 1)}$ and P_0 is the producer price. The variable s [0,1) is taken as our measure of the transport cost, instead of the less intuitive *r*. For a given producer price, *s* measures the percentage reduction in output (and thus profit) caused by losses in transit.

It is evident from (2) that the elasticity of demand is unaffected by the trade costs. Hence, the firm will charge a producer price that is a constant markup over marginal cost and is independent of trade costs. Specifically,

$$P_0 = \theta c \tag{3}$$

This mark-up is the same for each country, as preferences are identical.

The operating profit derived from the production of $X(s,t)$ is therefore

$$\tilde{\pi}(s,t) = (P_0 - c)X(s,t) = \beta(1+t)^{-\varepsilon}(1-s) \tag{4}$$

where $\beta = (\theta - 1)\theta^{-\varepsilon} c^{1-\varepsilon}$. Operating profit is a decreasing function of the transport cost and the tariff rate.

Let $v(s,t)$ denote the sum of consumer surplus and tariff revenue derived from the import of any variety subject to transport cost *s* and tariff *t*. Using (1), (2) and (3) gives

$$v(s,t) = \theta\beta(1+\varepsilon t)(1+t)^{-\varepsilon}(1\text{-}\ s) \tag{5}$$

This is also decreasing in *s* and *t*. The reason $v(s,t)$ is decreasing in the tariff is evident from equation (3). As the producer price is unaffected by trade costs, the imposition of a tariff results in no terms-of-trade improvement for the importing country. This is by design, for we wish to abstract from traditional optimal tariff considerations. Our focus is on the role of trade policy in enabling countries to avoid transport costs. Thus, in this model, the only way a country can benefit from a tariff is to induce foreign firms to engage in multinational production and thereby eliminate *s*.

2.2. Plant location

Basically, a firm's decision as to the number and location of its plants reflects a trade-off between trade costs and fixed costs. By locating a plant in a foreign country, the firm incurs *F* but avoids the trade costs associated with exporting to that market. Complicating this trade-off is the fact that the trade costs connected with exporting to any one market depend on the firm's plant-location choices in other markets, because the firm can export from any of its existing plants. Furthermore, with each new plant, the firm expands the number of sources from which it may export to any country.

We make two assumptions about plant location. First, if all countries choose zero tariffs, then all firms prefer single-plant production. A sufficient condition for this is $F \geq \beta(m-1)s_{ij}$ for all *i* and *j*. Second, we assume $\beta > F$, so that it is profitable for a firm to operate a foreign plant. This second assumption is sufficient to imply that there exists a high enough tariff, call it $\bar{t}$, that if any country were to impose $t \geq \bar{t}$ on an MFN basis, then all firms would locate a plant in that country, regardless of the policies of other countries. Such a tariff would maximize the country's consumer surplus, as all products would be locally produced.

3. The Two-Country Case

In this section we develop the simplest case, that of two countries, *A* and *B*. While this case does not directly address the issue of preferential trade agreements, it serves to illustrate the basic channel through which transport costs affect cooperation on trade policy. As a theory in itself, this model predicts that a decrease in transport costs, due to technological innovation for example, would tend to lower *overall* tariff levels. Whether or not this hypothesis is consistent with the actual time series is an open question.[11] The primary purpose, however, is to take an intermediate step toward establishing that geographical proximity facilitates trade policy cooperation.

With two countries, the plant location decision is binary: the firm either operates a single plant and exports to the other country, or it opens a plant abroad and exports nothing. As all firms in the same country are identical, the total profit of a country facing a foreign tariff of t is given by

$$\Pi(s,t) = n\left\{\beta + \max\left[\tilde{\pi}(s,t), \beta - F\right]\right\} \tag{6}$$

The tariff $\bar{t}$ is determined by $\tilde{\pi}(s,\bar{t}) = \beta - F$, which is where (6) is minimized.

Let $V(s,t)$ be the total consumer surplus plus tariff revenue, summed over the $2n$ goods consumed, for a country that imposes the tariff t. Using (5) and $\bar{t}$ gives,

$$V(s,t) = \begin{cases} n\left[\beta\theta + v(s,t)\right] & \text{for } t < \bar{t} \\ 2n\beta\theta & \text{for } t \geq \bar{t} \end{cases} \tag{7}$$

Combining (6) and (7), the total welfare of a country imposing tariff t and facing a foreign tariff t' is $W(s,t,t') = V(s,t) + \Pi(s,t')$.

Figures 1 and 2 illustrate components of welfare as functions of the tariff. Figure 1 shows total profit as a function of t for two different levels of s ($s < s'$). Total profit declines as the tariff level increases, for $t < \bar{t}$. At $\bar{t}$, the firms switch to multinational production and the tariff has no further effect on profits. Increasing the transport cost reduces the profit of a single-plant firm. This makes firms willing to switch to multinational production for a lower tariff. Thus, the critical tariff $\bar{t}$ falls to $\bar{t}'$.

---

[11] Hummels (1998) has called into question the conventional wisdom that transportation costs have fallen considerably in the post-war period. However, empirical work in this area is still in its infancy.

FIGURE 1

($s < s'$)

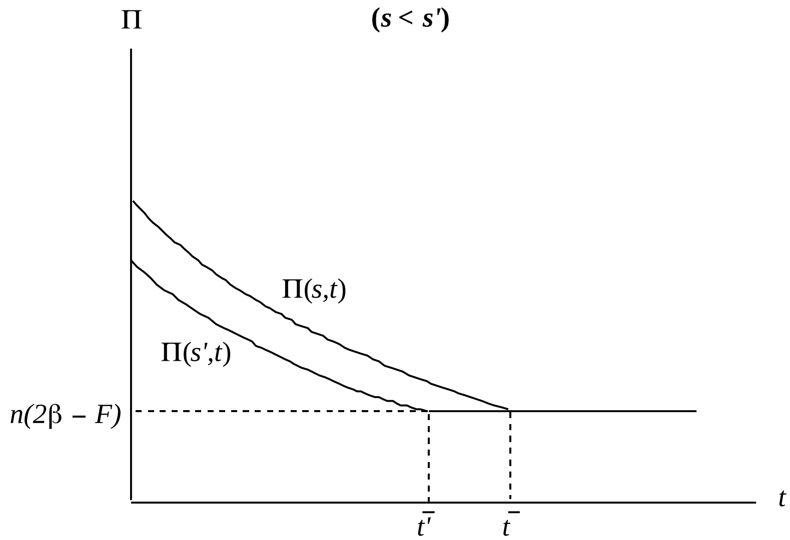
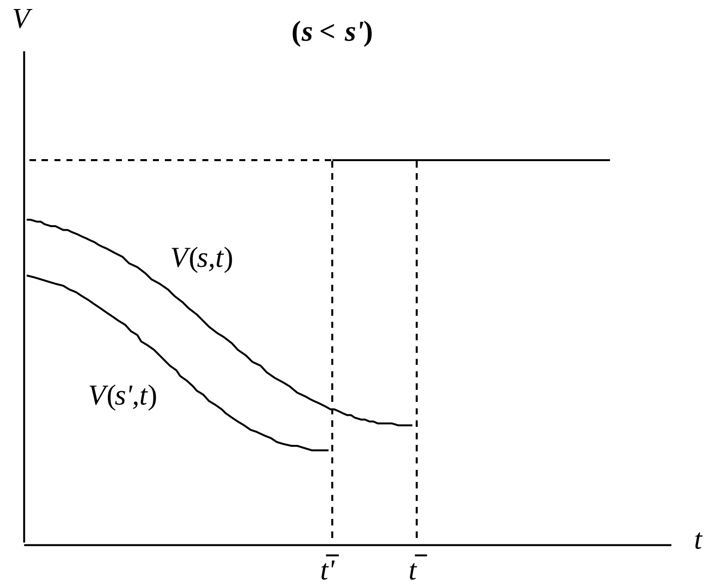


Figure 2 shows $V(s,t)$ as a function of t for the same two levels of s. $V(s,t)$ reaches a local maximum at zero and is declining in t, for all t below $\bar{t}$. Once a country's tariff becomes greater than $\bar{t}$, firms quit exporting to that country and invest in plants to produce locally. This benefits the country's consumers as the consumer price of the affected varieties falls from $(1+t)r^{-1}P_0$ to P_0. As long as the transport cost is positive, the consumers prefer the multinational outcome to free trade. Increasing the transport cost also reduces $V(s,t)$ for $t < \bar{t}$. This makes the incentive for a country to attract foreign investment even greater.

The fact that both $V(s,t)$ and $\Pi(s,t)$ are decreasing in t for $t < \bar{t}$ means that neither country gains from the imposition of a tariff, unless that tariff is high enough to induce multinational production. At that point the tariff is a beggar-thy-neighbor policy, with home consumers gaining at the expense of foreign firms.[12]

[12] As in Bagwell and Staiger (1999), the basis for cooperation is the fact a country can shift the cost of its tariff onto its neighbor. Here this cost is the fixed cost of a new plant, while in their model it is a terms of trade effect.

FIGURE 2
($s < s'$)



Whether free trade with single-plant firms or no trade with multi-plant firms is better in terms of overall world welfare depends upon parameters. As our interest is in free trade agreements, we assume free trade to be globally efficient, or $W(s,0,0) > W(s,\bar{t},\bar{t})$. Using (6) and (7), the condition for the efficiency of free trade can be written as

$$\frac{F}{\beta} > s(\theta + 1) \tag{8}$$

The left-hand side of (8) is less than 1, otherwise the firms would never find it profitable to establish a second plant. Subject to that, it is clear that the fixed cost of setting up the second plant must be suitably large, or the transport cost suitably small for this condition to hold.

3.1. Self-enforcing trade agreements

In this section, we consider an infinitely repeated game based on the model developed in the previous subsection. Suppose there are an infinite number of discrete time periods, each consisting of a game in which governments choose tariffs, and firms make location and production decisions, in that order.[13] A strategy is now an infinite sequence of functions, one for each period, mapping the history of past play into current actions. Agents discount payoffs in future periods by a factor $\delta < 1$.

In repeated games such as this it is common to find a large set of subgame perfect equilibria. As the purpose of this paper is to explore the limits of trade policy cooperation, however, we shall be interested only in the Pareto frontier of this set, defined with respect to the total welfare levels of the two countries. To construct the set of Pareto efficient (PE) subgame perfect equilibria, we consider first the one-shot version of the game.

In the one-shot game, each government's dominant strategy is to impose a tariff high enough to induce every foreign firm to set up a second plant. Thus, the one-shot equilibrium involves multinational production by all firms and tariffs no less than $\bar{t}$ in each country. Moreover, the welfare received in this equilibrium is the lowest to which any country can be held in any equilibrium. Thus, Nash reversion is an optimal penal code (Abreu, 1988). It follows that to construct the complete set of subgame perfect equilibria, one need only look for tariff pairs (t_A, t_B) supportable by the threat that, if any player should deviate, then the one-shot equilibrium would prevail in all future periods.

The tariff pair (t_A, t_B) can be supported as a subgame perfect equilibrium if,

$$(1-\delta)\left[W(\bar{t},t_j)-W(t_i,t_j)\right]\leq\delta\left[W(t_i,t_j)-W(\bar{t},\bar{t})\right],\quad i,j=A,B,\ i\neq j \tag{9}$$

[13] The assumed timing of the one-shot game (specifically, that firms take tariffs as given when making location decisions) is fairly standard in the literature but not universal. In Brander and Spencer (1987), for example, the location decision occurs first. Time consistency is the usual issue raised in this context; however, time consistency is less of a problem in a repeated game. Indeed, there will be generally a host of different equilibria, all of which are time consistent, regardless of the timing of the stage game. The virtue of assuming governments move first within the stage game is that it implies that for certain discount factors, free trade will not be sustainable. Thus, it is an appropriate model for studying the limits of trade policy cooperation. Whether it is more realistic to use one timing assumption or the other is ultimately an empirical question. This was addressed somewhat in the introduction, but a thorough treatment is beyond the scope of this paper.

where the parameter s is suppressed for notation brevity. The left-hand side of condition (9) is country i's temptation to deviate from the equilibrium. It is the one-shot gain from choosing an investment-inducing tariff instead of t_i. The right-hand side is the enforcement. It is loss from switching to the punishment equilibrium in the next and all future periods relative to remaining at (t_A, t_B).

It is useful to decompose (9) into consumer and producer components as follows:

$$V(\bar{t}) - V(t_i) \leq \delta\left[\Pi(t_j) - \Pi(\bar{t})\right] \quad (10)$$

In other words, the gain to consumers associated with multinational production must be less than the discounted loss in firm profits. Recall that the gain to consumers from multinational production (the left-hand side of (10)) is an increasing function of both the tariff rate and the transport cost. The loss to producers (the right-hand side of (10)) is a decreasing function of both the tariff and the transport cost. This has two important implications:

Proposition 1: If any tariffs less than $\bar{t}$ can be supported in equilibrium, then bilateral free trade can also be supported. Thus, zero is the unique PE equilibrium tariff for both countries. Conversely, if free trade is not an equilibrium, then the PE equilibrium tariff is $\bar{t}$ or greater.

The proof of proposition 1 is straightforward. There can be no equilibrium in which one tariff is $\bar{t}$ and the other is less than $\bar{t}$, because this would give the low tariff country a payoff less than $W(\bar{t},\bar{t})$. For tariffs less than $\bar{t}$ in both countries, bilateral free trade gives the maximum enforcement, the minimum temptation, and the highest welfare. Thus, if any $t < \bar{t}$ is an equilibrium, free trade is the PE equilibrium. If free trade is not an equilibrium, then we are left with only the one-shot equilibrium (for $t \geq \bar{t}$ both enforcement and temptation are zero).

Proposition 2: The lower the transport cost the larger is the set of discount factors that will support free trade as an equilibrium.

This is because lowering the transport cost increases enforcement and decreases temptation. To be more precise, we substitute equations (6) and (7) into condition (10) and set $t_i = t_j = 0$, yielding the following condition for a free trade equilibrium:

$$\frac{F}{\beta} \geq s\left(\frac{\theta}{\delta} + 1\right) \tag{11}$$

The only difference between this and condition (8) (the condition for the efficiency of free trade) is the presence of the discount factor in the right-hand side of (11). From (11) it follows that a larger fixed cost, a higher discount factor or a lower transport cost all favor free trade agreements.

The above results are illustrated in figures 3 and 4. In figure 3, the two sides of condition (10) are drawn separately, as functions of t, for two different levels of s. The shaded area represents the excess of enforcement over temptation. For transport cost s, the shaded region corresponds to an interval of tariffs including zero. Thus, free trade is the PE equilibrium tariff. For the higher transport cost s', free trade is not an equilibrium, as temptation exceeds enforcement for all $t < \bar{t}'$. In this case, the PE equilibrium tariff is $\bar{t}'$. It was noted earlier that $\bar{t}$ is decreasing in s implying that for even higher transport costs (higher than s') the PE equilibrium tariff would be lower. Thus, as the transport cost falls, the PE equilibrium tariff rises until the transport cost reaches a critical level (where (10) is satisfied with equality) at which point the tariff drops to zero and remains there for all lower transport costs.

FIGURE 3

($s < s'$)

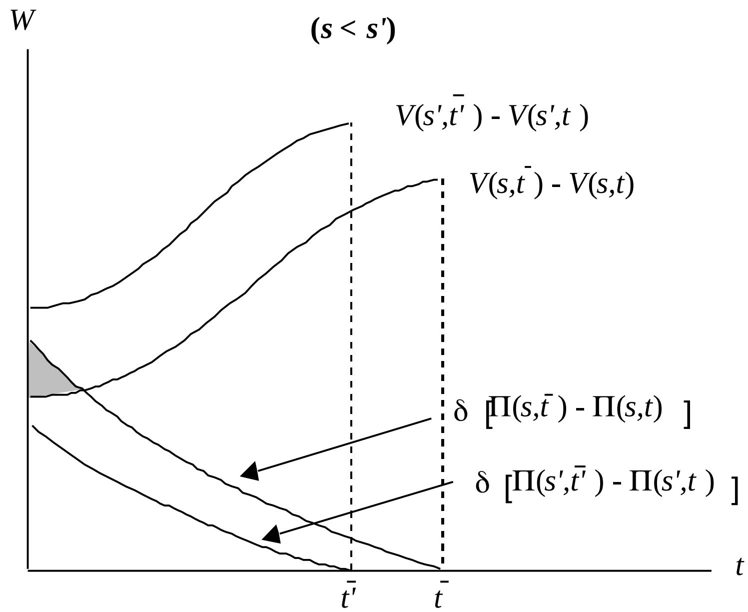
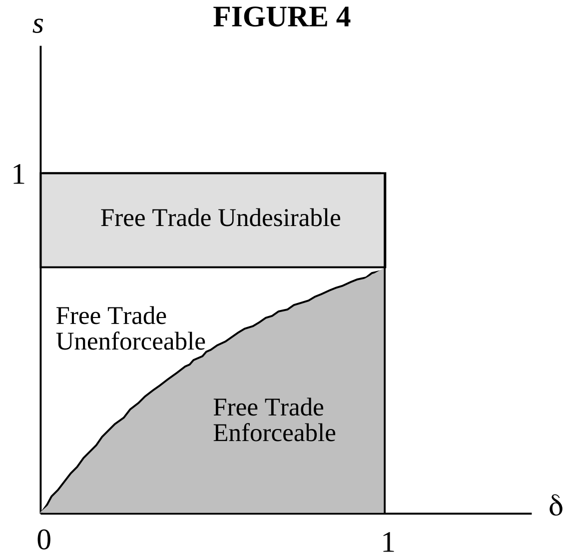


Figure 4 depicts the equilibria corresponding to different combinations of the discount factor and the transport cost. Above a certain level of *s*, condition (8) is violated and the countries would be better off in the punishment equilibrium. Below this level of *s*, the free trade is enforceable for a high enough discount factor. The minimum discount factor necessary to support free trade increases with *s*.

FIGURE 4



3.2. Robustness to alternative models of cooperation

The inverse relationship between cooperation and the transport cost is unambiguous in our model because a reduction in *s* lowers temptation and raises enforcement. The fact that both of these components move in the "right" direction suggests that our conclusion may be robust to alternative models of trade cooperation. The repeated games approach to trade agreements used here has been criticized from at least two different directions. On the one hand, the temptation to defect stems from the arguably strong assumption that punishment is delayed long enough for the private sector respond and generate substantial payoffs to the defecting country. One might question why the punisher would wait so long. In defense of this assumption, Staiger (1995) points out that, in practice, punishments typically *are* delayed many months. On the other hand, enforcement stems from the assumption that countries are willing to revert to a Pareto inefficient punishment equilibrium, even though they may prefer to renegotiate instead. This issue has been addressed in the literature, and space prevents a thorough treatment of the alternatives (again see, Staiger, 1995). The point is, however, whatever model of cooperation one uses, as long as the likelihood of cooperation is positively related to the benefits of cooperation and/or negatively related to the incentive for unilateral trade restriction, the conclusion that low transport costs foster cooperation should go through.

3.3. Robustness to sunk costs and industry growth

Thus far we have assumed there is a fixed cost of operating an MNE, which is paid each period. Examples might include administrative expenses, interest on capital, maintenance of equipment, or rent on foreign land. One might also suppose that there are sunk costs associated with setting up MNE—once-and-for-all costs incurred when the firm opens its foreign plant but irrecoverable if the plant subsequently closes.[14] In this section, we show that our results are robust to the addition of sunk costs to our model. Moreover, the model can be modified to accommodate even the absence of fixed costs (i.e., $F = 0$). One such modification, allowing the industry to grow over time, is demonstrated explicitly, and the results are found to be substantially the same.

The presence of sunk costs changes the plant location problem, because the cost of opening a new plant is now greater than the cost of maintaining an existing one. This implies that the tariff necessary to induce firms to switch to multi-plant production is higher than the tariff necessary to prevent them from reverting to single-plant production. Specifically, suppose the firm pays a one-period sunk cost G to set up its foreign plant and a fixed cost F in each period of the plant's operation. The minimum tariff necessary to induce a single-plant firm to build a second plant becomes $\hat{t}$, defined by $\tilde{\pi}(s,\hat{t}) = \beta - (1-\delta)G - F$,[15] whereas the tariff necessary to prevent a multi-plant firm from shutting down its foreign plant continues to be $\bar{t}$, defined by $\tilde{\pi}(s,\bar{t}) = \beta - F$. Evidently, $\hat{t} > \bar{t}$.

Now consider the effect of sunk costs on trade agreements. The two countries will be able to sustain to a free trade agreement if and only if,[16]

$$\frac{(1-\delta)G + F}{\beta} \geq s\left(\frac{\theta}{\delta} + 1\right) \tag{12}$$

If (12) fails, then a somewhat interesting time path may emerge. In the first period, the countries impose tariffs large enough to induce all firms go multinational ($t_1 \geq \hat{t}$), while in the

---

[14] If the set-up cost is recoverable, then it is equivalent to a fixed cost, as firm's decisions to start and stop multinational production are symmetric. Introduction of this type of set-up cost would occasion no change in the model or its results.

[15] The firm's decision to open a second plant is now influenced by the entire sequence of future tariffs. If the sequence is constant or increasing on average, then $\hat{t}$ is sufficient to induce multi-plant production. If the sequence were declining, then a tariff higher than $\hat{t}$ would be required.

[16] This is but a slight modification of (11).

second, and all future periods, the countries may agree to set tariffs as low as $\bar{t}$. Such an agreement has no effect on plant location or welfare (as there is no trade), but it is an equilibrium nonetheless. Apart from this time path, however, our results are unchanged: tariffs are either zero or (at least) $\bar{t}$ depending on whether or not *s* is low enough to satisfy (12).

While the addition of sunk costs to our model has little effect, the elimination of fixed costs creates a problem. We can no longer maintain the assumption that free trade induces single-plant production (or that free trade with single-plant production is efficient), once all firms have established a second plant. Moreover, $\bar{t} \leq 0$, and thus, after the first period, free trade may obtain whether (12) holds or not.

The basic problem with the zero-fixed-cost case is that once all firms have established foreign plants in period 1, no firm ever has to pay the sunk cost again. If the sunk cost were to recur periodically, or if new (single-plant) firms were to enter the market over time, then this problem would evaporate. Indeed, so long as in every period there are least two firms (one in each country) faced with choice between single-plant production and paying the sunk cost, governments must maintain their tariffs at $\hat{t}$ or higher to induce those firms to go multinational. Thus, trade agreements involving $\bar{t}$ would cease to be PE equilibria.

To illustrate this formally, suppose the number of firms grows at a constant rate *g*, and that all firms begin with a single plant. In each period τ, there are at least $gn_{\tau-1}$ firms that must decide whether or not to become multinationals, where $n_{\tau-1}$ is the total number of firms in each country at $\tau - 1$. The minimum tariff necessary to induce multi-plant production is $\hat{t}$. While this tariff is irrelevant to all firms that have already sunk the cost of multi-plant production in earlier periods, it influences each new cohort of firms.

Let $\gamma \equiv (1+g)^{-1}$. To ensure that welfare is bounded, we require the growth rate to be less than the discount rate, or $\gamma > \delta$. Let the state variable σ_τ denote the share of single-plant firms in the population at the beginning of period τ (i.e., before the plant location decision). The state evolves as follows: if t_τ induces multi-plant production, then $\sigma_{\tau+1} = 1-\gamma$; otherwise, $\sigma_{\tau+1} = 1-\gamma + \sigma_\tau\gamma$. There are two possible steady states, $\sigma = 1-\gamma$ and $\sigma = 1$.

Suppose the two countries are party to a free trade agreement that they expect will last forever. The temptation to deviate is $T(\sigma_\tau, s) = (1-\delta)\beta\theta s\sigma_\tau$, which is the short-term gain a country obtains from inducing all existing single-plant foreign firms to go multinational in period τ. This is weighed against enforcement, $E(\sigma_\tau, s) = \delta\left(\frac{1-\gamma}{\gamma-\delta} + \sigma_\tau\right)\left[(1-\delta)G - \beta(\theta+1)s\right]$, which is the long-term welfare loss from reversion to the punishment equilibrium, discounted one period. To understand $E(\sigma_\tau, s)$, note that each firm that goes multinational causes a welfare loss of $(1-\delta)G - \beta(\theta+1)s$. This is multiplied by the share of single-plant firms present at time τ, plus a term representing future entrants. A long-term free trade agreement is sustainable in period τ, if and only if, $E(\sigma_\tau, s) \geq T(\sigma_\tau, s)$.

A free trade agreement causes σ to rise over time and eventually converge to one. In the free trade steady state, therefore, we must have $E(1,s) \geq T(1, s)$, or

$$\frac{(1-\delta)G}{\beta} \geq s\left(\frac{\gamma\theta}{\delta} + 1\right) \tag{13}$$

Let $\bar{s}$ denote the value s that satisfies (13) with equality. Also, define $\hat{s}$ such that $E(1-\gamma, \hat{s}) = T(1-\gamma, \hat{s})$. Note that $0 < \bar{s} < \hat{s} < 1$.

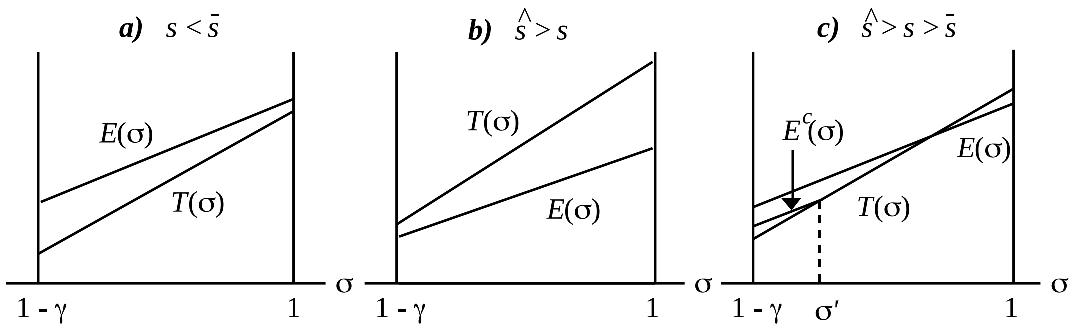
Proposition 3: Bilateral free trade in every period is the unique PE equilibrium, if and only if $s \leq \bar{s}$. If $s \geq \hat{s}$, then $t_\tau \geq \hat{t}$ for all τ. If $\bar{s} < s < \hat{s}$, then either $t_\tau \geq \hat{t}$ for all τ, or there is an equilibrium cycle (alternating periods of free trade and high tariffs).

Proof in appendix.

Proposition 3 is illustrated in figure 5. If $s \leq \bar{s}$, then enforcement everywhere exceeds temptation. Thus free trade is sustainable at every step along the path to the steady state. If $s \geq \hat{s}$, then temptation is too great to sustain free trade (or any other agreement) in any state. If $\bar{s} < s < \hat{s}$, then the temptation and enforcement schedules cross, and free trade cannot be sustained in the long term. At best, there may be a cycle, where free trade occurs for a few periods and then collapses. The enforcement associated with such a cycle ($E^c(\sigma)$ in figure 5c) is everywhere less than $E(\sigma)$. It may also be everywhere less than $T(\sigma)$, in which case $t_\tau \geq \hat{t}$ for all τ. However, if $E^c(1-\gamma) > T(1-\gamma)$ (as illustrated), then free trade can persist until the share of single-plant firms reaches σ', at which point the tariff jumps for one period and the cycle repeats. In general, the gap between $\hat{s}$ and $\bar{s}$ is pretty small (ranging from 0 to 0.1) and the set on which cycles occur is considerably smaller (if not null).

Clearly, the dynamic nature of the problem adds some complexity, most notably the possibility of cycles. However, the salient result, that low transport costs give rise to free trade and high transport costs do not, continues to go through. Thus neither the presence of sunk costs nor the absence of fixed costs undermines our basic conclusion.

FIGURE 5



4. A Triangular Model

Returning to our original model with fixed costs, we now introduce a third country *C*, which is identical to *A* and *B* in every respect except location. The transport cost between *A* and *B* is *s*, while the transport cost between *A* and *C* and between *B* and *C* is assumed to be $s' > s$. One can think of the three countries as being located at the corners of an isosceles triangle, with *C* being geographically remote but equidistant from its two trading partners. The purpose of this model is to determine whether a preferential trade agreement between neighbors, *A* and *B*, is in some sense more likely than one between separated countries like *A* and *C*. As part of this, we will need to establish that PTAs can occur despite the presence of alternatives such a global free trade.

There are three preliminary issues that must be dealt with before we begin the analysis. The first is a tie-breaking rule for when firms are indifferent over the locations of their foreign plants. In a three-country world, a firm must choose whether to operate one, two or three plants and, if it operates two plants, where to locate the second plant and whether to export from there or from home. The optimal choice is usually determinate. However, if a firm is headquartered outside of a bilateral FTA, decides to locate a plant within the FTA, and supply both FTA markets from that plant, it will be indifferent as to where in the FTA to locate. In the aggregate, these choices matter because they affect the distribution of consumer surplus within the FTA. The FTA country receiving the smaller investment share will have a correspondingly smaller share of FTA consumer surplus and greater tendency to deviate from the FTA. The most resilient FTA therefore will be one in which the investment shares are equal. For this reason and to preserve as much symmetry in the model as possible, we assume that, when outside firms are indifferent, half of them locate in one FTA country and half in the other.

The second issue is that the added dimensionality has the potential to produce a large set of Pareto efficient equilibria. To limit this complexity (which, combined with added complexity of firm location choice, becomes intractable) and focus on outcomes that resemble actual trade agreements, we restrict attention to equilibria that satisfy a mild form of reciprocity, namely, if $t_{ij} < \bar{t}$ then $t_{ji} < \bar{t}$ (for all $i, j = A, B, C, i \neq j$). That is, a country only reduces its tariff if it receives some tariff reduction from its trade partner in return. This is certainly a conventional if not entirely reasonable requirement for a trade agreement to satisfy.

Finally, we note one important similarity between the two and three country models. Every tariff in a PE equilibrium must be either zero or prohibitive. This was established in Proposition 1 for the two-country case. For the triangular model we have the following lemma:

Lemma 1: In any PE equilibrium, if $t_{ij} > 0$, then exports of X from country j to country i must be zero (for all $i, j = A, B, C$).

Proof in appendix.

The implication of this lemma is that we can restrict attention to equilibria involving only two tariff levels: zero and $\bar{t}$. The intuition is that tariffs are of no benefit to producers or consumers in this model, unless they affect plant location. Thus, an equilibrium in which the tariff matrix contains only zeros and $\bar{t}$ s must Pareto dominate any other equilibrium that supports the same configuration of plants. It turns out that there are no equilibrium plant configurations that cannot be reproduced by a matrix of zeros and $\bar{t}$ s. It should be noted that lemma 1 does not depend on reciprocity.

Lemma 1 implies a dramatic reduction in the search for equilibria. In fact, there now are only six cases to be considered, which can be organized into three categories: MFN equilibria; exclusive bilateral agreements; and hub-and-spoke arrangements.

4.1. MFN equilibria

4.1.1 The one-shot Nash equilibrium

The tariff $\bar{t}$ continues to be determined by $\tilde{\pi}(s,\bar{t}) = \beta - F$. This is the smallest MFN tariff country *B* would need to impose to induce a firm operating in country *A* to open a plant in *B*, given any tariff imposed by *C*. Firms not operating in *A* would be even more inclined open a plant in *B*, as $s' > s$. By imposing $\bar{t}$, a country maximizes its consumer surplus (as all goods come to be produced locally), regardless of the behavior of other governments. Thus, this continues to be the dominant strategy for all countries in the one-shot game. The one-shot equilibrium therefore gives the following payoffs:

$$\bar{V} = 3n\theta\beta \quad \text{and} \quad \overline{\Pi} = n(3\beta - 2F) \tag{14}$$

for all countries. $\bar{V}$ is the surplus from consuming 3*n* goods subject to zero trade costs, while $\overline{\Pi}$ is the profit of a country's *n* firms each operating three plants. $\bar{V} + \overline{\Pi}$ is the lowest welfare any country can experience in any subgame perfect equilibrium, and thus it continues to be the severest possible punishment.

4.1.2 Global free trade (FT)

With no tariffs, all firms prefer single plant production and export everything from their home plants. A firm in country *A* sells to the *A* market without trade cost, to the *B* market subject to *s*, and to the *C* market subject to s', yielding a total profit of $\tilde{\pi}(0,0) + \tilde{\pi}(s,0) + \tilde{\pi}(s',0) = \beta(3 - s - s')$. Firms in *B* receive the same. A firm in *C* sells to both its export markets at s', yielding $\tilde{\pi}(0,0) + 2\tilde{\pi}(s',0) = \beta(3 - 2s')$. Combining these profits with the corresponding consumer surpluses produces the following payoffs:

$$V_A^{FT} = V_B^{FT} = n\theta\beta(3 - s - s') \text{ and } \Pi_A^{FT} = \Pi_B^{FT} = n\beta(3 - s - s') \qquad (15)$$

$$V_C^{FT} = n\theta\beta(3 - 2s') \text{ and } \Pi_C^{FT} = n\beta(3 - 2s') \qquad (16)$$

Both producers and consumers in *A* and *B* are better off in *FT* than are their counterparts in country *C*, due to the assumed geography. Thus, *C* is the country most likely to deviate from a global free trade agreement. Using (14) and (16) in (10), the condition for global free trade to be subgame perfect is $\bar{V} - V_C^{FT} \le \delta(\Pi_C^{FT} - \bar{\Pi})$, or

$$\frac{F}{\beta} \ge s'\left(\frac{\theta}{\delta} + 1\right) \qquad (17)$$

It follows that if country *C* is too far away from *A* and *B*, global free trade will not occur, even if it is Pareto efficient.

4.2. Exclusive bilateral agreements

4.2.1 FTA between adjacent countries (AB)

With free trade between *A* and *B*, and all other tariffs equal to $\bar{t}$, a firm headquartered in *A* or *B* will open a second plant in *C* but continue to supply its FTA partner from its home plant. This yields a profit of $2\tilde{\pi}(0,0) - F + \tilde{\pi}(s,0) = \beta(3 - s) - F$. A firm headquartered in *C* opens a second plant either in *A* or *B* and exports from there to the other FTA member. Thus, regardless of where the C firm locates its second plant, it receives the same profit as do the A and B firms.

Using the assumption that half of C's foreign plants locate in each FTA member, the payoffs associated with an FTA between A and B are

$$V_A^{AB} = V_B^{AB} = n\theta\beta\left(3 - \tfrac{3}{2}s\right) \text{ and } \Pi_A^{AB} = \Pi_B^{AB} = n\beta(3 - s) - F \tag{18}$$

$$V_C^{AB} = \bar{V} \text{ and } \Pi_C^{AB} = n\beta(3 - s) - F \tag{19}$$

Country C is actually better off in the presence of AB than are its members. This is because C suffers no transport costs on its imports yet benefits from the low trade costs between A and B on its exports. Thus, A and B are the countries most likely to deviate. Using (18) in (10), AB can be supported as an equilibrium if and only if,

$$\frac{F}{\beta} \geq s\left(\frac{3}{2}\frac{\theta}{\delta} + 1\right) \tag{20}$$

Thus, an FTA between the adjacent countries is a candidate for equilibrium if the transport cost between A and B is small enough. We shall postpone discussing whether this is a Pareto efficient equilibrium until after we have examined the other cases.

4.2.2 FTA between distant countries (AC)

This case is the same as AB except that the relevant transport cost is s' instead of s. Thus, the payoffs to an FTA between A and C are

$$V_A^{AC} = V_C^{AC} = n\theta\beta\left(3 - \tfrac{3}{2}s'\right) \text{ and } \Pi_A^{AC} = \Pi_C^{AC} = n\beta(3 - s') - F \tag{21}$$

$$V_B^{AC} = \bar{V} \text{ and } \Pi_B^{AC} = n\beta(3 - s') - F \tag{22}$$

and this can be supported as an equilibrium if and only if,

$$\frac{F}{\beta} \geq s'\left(\frac{3}{2}\frac{\theta}{\delta} + 1\right) \tag{23}$$

Comparing condition (23) with (17) it is evident that if global free trade fails to be an equilibrium than so too does the FTA between distant countries.

4.3. Hub and spoke arrangements

If bilateral trade agreements were restricted to being customs unions, rather than free trade areas (the difference being that customs unions require a common external tariff), then the foregoing cases would be exhaustive. However, if members of a bilateral agreement are permitted to have different external tariffs, then it is possible for one member of an FTA to sign a second FTA without including its partner from the first. This is known as a hub and spoke arrangement[17]—the country that is a member of both agreements is the hub, while the other two countries are spokes. There are two generic cases of this.

4.3.1 Country A as the hub (AH)

If country *A* is the hub, firms headquartered in *A* enjoy tariff-free access to *B* and *C* and thus receive the same profit as under *FT*. Firms headquartered in *B* locate a plant in *C* in order to serve that market, and export to *A* from their home plants. This produces the same profit as under *AB*. A firm headquartered in *C* locates a plant in either *A* or *B* and exports from there, also generating the same profit as under *AB*. The payoffs are therefore,

$$V_A^{AH} = V_A^{AB} \text{ and } \Pi_A^{AH} = \Pi_A^{FT}$$

$$V_B^{AH} = V_B^{AB} \text{ and } \Pi_B^{AH} = \Pi_B^{AB}$$

$$V_C^{AH} = n\theta\beta\left(3 - s'\right) \text{ and } \Pi_C^{AH} = \Pi_C^{AB}$$

The only difference between this and *AB* is that now *A*'s profit is greater and *C*'s consumer surplus is lower. In other words, *A* saves from having to invest in *C* at the expense *C*'s

[17] See Wonnacott (1990) uses this terminology.

consumers. One implication of this is that if both *AB* and *AH* are equilibria, then neither one Pareto dominates the other.

Whether *B* or *C* is the country more likely to deviate from *AH* depends on the transport costs. For high values of s', consumer surplus in *C* will be low and thus *C* will be the more likely to deviate. Otherwise, *B* is the potential deviator, and since *B* receives the same payoff in *AH* as in *AB*, the existence condition for this equilibrium is the same for *AB*. The condition for the existence of *AH* is therefore

$$\frac{F}{\beta} \geq \max\left[s\left(\frac{3}{2}\frac{\theta}{\delta}+1\right), s'\left(\frac{\theta}{\delta}+s\right)\right] \tag{24}$$

4.3.2 Country C as the hub (CH)

Finally, there is the possibility that country *C* is the hub, maintaining bilateral free trade agreements with *A* and *B*. Firms headquartered in *C* receive the same profit as under *FT*. Firms headquartered in *B* locate a plant in either *A* or *C* and export from there, generating the same profit as under *AC*. Firms headquartered in *A* behave symmetrically and thus receive the same profit as firms headquartered in *B*. Thus, the payoffs to *CH* are

$$V_A^{CH} = V_B^{CH} = V_A^{AC} \quad \text{and} \quad \Pi_A^{CH} = \Pi_B^{CH} = \Pi_A^{AC}$$

$$V_C^{CH} = V_C^{AH} \quad \text{and} \quad \Pi_C^{CH} = \Pi_C^{FT}$$

and existence condition for this equilibrium is identical to that of *AC*. As with the previous hub-and-spoke arrangement, if both *AC* and *CH* are equilibria, then neither one Pareto dominates the other.

Before moving on to the results is should be noted that all of the equilibrium conditions derived above can be adapted to include sunk costs. This would amount to adding $(1-\delta)G$ to the numerator in the left-hand side of (17), (20), (23) and (24). Using the dynamic, zero-fixed-cost specification, if we restrict attention to steady-state equilibria, the equilibrium conditions can be found by replacing θ with $\gamma\theta$ in the right-hand side of (17), (20), (23) and (24). If we allow cycling, the problem becomes substantially less tractable.

4.4 Results

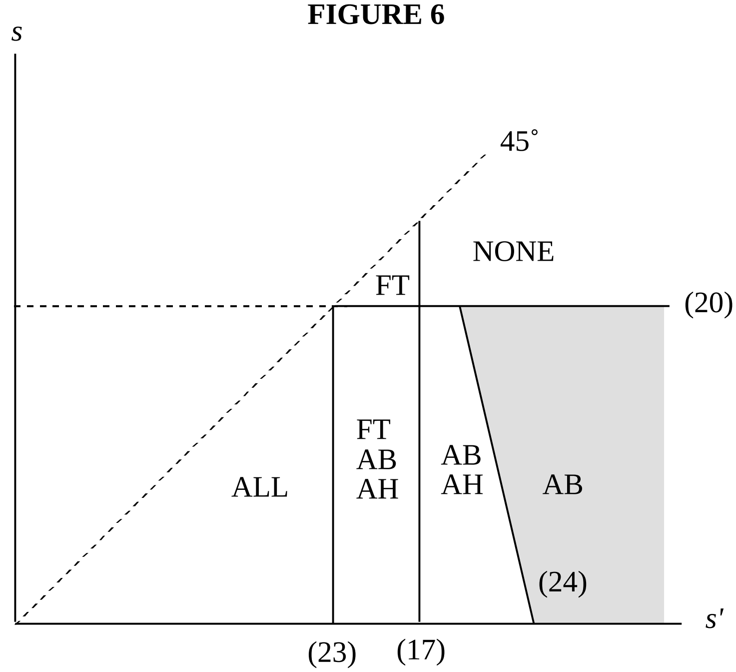
A comparison of the existence conditions for the six cases described above is found in figure 6. Figure 6 shows the different trade agreements that exist for every combination of s and s' such that $s < s'$. The boundaries separating the various sets are the combinations of s and s' that satisfy the equilibrium conditions with equality, and are labeled according the corresponding equation numbers.

One conclusion that can be drawn immediately from figure 6 is that if s is below the threshold (20) that permits a preferential agreement between A and B, then an exclusive bilateral *FT*A between A and B is the only equilibrium (other than the punishment equilibrium) to survive a high enough s'. The shaded region in figure 6 shows this. The next proposition states the result formally.

Proposition 4: For all $s \leq \frac{F}{\beta}\left(\frac{2\delta}{2\delta + 3\theta}\right)$ and $s' > \left(\frac{F}{\beta} - s\right)\frac{\delta}{\theta}$, *AB* is the unique PE equilibrium.

Proposition 4 is the central result of the paper. It establishes that if two countries have low transport costs between them and high transport costs in relation to the rest of the world, then those two countries must have an exclusive bilateral FTA. This is strongly indicative of a regional bias in PTAs.

FIGURE 6



If s is above the threshold (20), then either *FT* is the outcome (for low s') or no cooperation is possible (for high s'). This is the same as in the two-country model. For all other transport costs, there are multiple trade agreements that satisfy the subgame perfection requirement, and thus we must apply the Pareto criterion to refine the equilibrium set any further. We have already noted that when both *AB* and *AH* exist, neither Pareto dominates the other, so we focus the remainder of this section on cases in which (17) and (20) are both satisfied [i.e., points in Figure 6 below (20) and left of (17)].

Proposition 5: If (17) and (20) are both satisfied, then

a) *FT* and *AH* are PE equilibria

b) *AB* is a PE equilibrium if and only if, $s \le 2(\theta + 1)s' - (F/\beta)$.

Proof in appendix.

Proposition 5b implies that it is possible for *AB* to be ruled out on grounds of Pareto efficiency. This will occur whenever *C* prefers *FT* to *AB* (the condition stated in proposition

5b is the converse). If the cost of shipping goods from *C* is low relative to the fixed cost of setting up a new plant, then *C*'s firms benefit from global free trade (relative to *AB*) by more than *C*'s consumers benefit (relative to *FT*) from having foreign goods locally produced. Thus, a low s' relative to *F* and *s* will tend to promote global free trade.

Finally, consider the only two cases (other than the punishment equilibrium) in which countries *A* and *B* do not have free trade between them. These are *AC* and *CH*. Condition (23) must be satisfied for these equilibria to exist [i.e., points to the left of (23) in figure 6]. The next proposition states the conditions under which they are Pareto efficient.

Proposition 6: If (23) is satisfied, then

a) *AC* is a PE equilibrium if and only if, $s \geq \frac{(F/\beta) - \theta s'}{1+\theta}$

b) *CH* is a PE equilibrium if and only if, $s \geq (\theta + 2)s' - (F/\beta)$

c) If $\delta \leq 3/4$, then *CH* is a PE equilibrium and *AC* is not.

Proof in appendix.

Proposition 6 suggests that *AC* is an unlikely equilibrium. It is ruled out for discount factors less than three-fourths (6c). When $\delta > 3/4$, *AC* can only survive for high values of *s* and s', corresponding to the upper corner of the triangular region of figure 6 in which all agreements are PE equilibria. The condition in proposition 6a is the condition for country *B* to prefer *AC* to *FT*.

The hub and spoke arrangement *CH* is more robust. When it exists it is likely to be Pareto efficient, because the hub country *C* enjoys free trade for its exports to both *A* and *B*, while receiving investment from those *A* and *B* firms that locate in *C* so as to export to the other spoke.

The main conclusion to be drawn from propositions 5 and 6 is that if transport costs are fairly low across the board, then we can expect to see agreements involving country *C*. These will typically be *in addition to* free trade between *A* and *B*, as in the cases of *FT* and *AH*. But they may occur instead of free trade between *A* and *B*, as in the cases of *AC* and *CH*. Thus, it is not necessarily the case that, if all three countries are relatively close to each other, free trade will occur between the two closest countries. Empirically, therefore, we should expect to see a slightly weaker correlation between tariffs and transport costs (or distance) within well-defined regions than between them.

4.5 Targeted punishment

The previous sections operated under the assumption that defection from an equilibrium would result in worldwide trade war. The value of this assumption was that it enabled us to find the largest possible set of supportable trade agreements. Despite its effectiveness, however, complete trade war is not necessarily the most reasonable punishment, particularly in cases where other, less draconian, punishments can produce efficient outcomes as well. If one country defects from a three-country trade agreement, it may be considered undesirable that the non-defecting countries should raise tariffs against each other for no other reason than to punish the defector. A reasonable alternative might be to suppose that the non-defecting countries target their tariff increases, applying $\bar{t}$ only to imports from the defecting country, while maintaining tariffs toward each other at pre-defection levels, provided this is equilibrium behavior.[18] If this is not equilibrium behavior, then countries revert to generalized trade war.

In the interest of space, we shall not present an analysis of the targeted punishment case, only the results. Interested readers are referred to Ludema (1999) for details. It turns out that many of the results from the generalized punishment case are preserved under targeted punishment: *AB* is the unique equilibrium for low *s* and high *s*′; *FT* is an equilibrium for low values of both *s* and *s*′, and if *s* and *s*′ are both high, no cooperation is possible. However, there are some significant differences as well. Under targeted punishment, we find: (1) hub-and-spoke arrangements are eliminated as equilibria; (2) the range of parameters for which global free trade can occur is reduced; and (3) the range of parameters for which *AB* is the unique equilibrium is substantially increased. Thus, targeted punishment significantly strengthens the regional bias of PTAs.

5. Conclusions

The strong empirical relationship that seems to exist between proximity and preferential trade is an issue worthy of the attention of economists interested in economic geography and commercial policy. It is not easily explained using standard trade theory, even when account is taken of trade agreement imperfections due to limited enforceability. This paper uses a simple model of horizontal MNEs, which when combined limited enforceability, produces a strong inverse relationship between physical transport costs and trade policy cooperation. In the three-country model, this produces a regional bias in PTAs.

This paper has highlighted the possible role of horizontal MNEs. However, it may be that multinational activity alone is too small in relation to total trade flows to account for the geography of PTAs. Thus, it would be useful to extend the analysis to other models of trade that exhibit the home market effect. Some progress has been made using the models of Venables (1985 and 1987). Though the analytics of these cases are rather untidy, numerical simulations have produced results similar to those presented here.

---

[18] This is what Maggi (1999) refers as a “multilateral enforcement mechanism.”

The simple MNE model used in this paper has the counter factual implication that trade and FDI are perfect substitutes. It is well known, however, countries with large bilateral FDI flows also tend to have large trade bilateral trade flows. While the literature MNEs has not completely settled on an explanation for this, it is clear that to account for it there must be "mixed" equilibria, in which both single-plant and multi-plant firms operate simultaneously. This possibility arises in most of leading, though more complicated, MNE models. Thus, our results should be reconsidered using a model such as Horstmann and Markusen (1992).

Another worthwhile next step in this research would be to increase the number of countries. Our results seem to suggest that trade blocs may come in clusters, and there may be a maximum geographic size of an bloc. It is likely that this maximum size would grow as the transport cost falls

Despite its limitations, this paper is a step toward understanding the geography of trade agreements. The model features transport costs, economies of scale, and imperfect competition — the same recipe that has proven successful in the economic geography literature in explaining a wide variety of geographic phenomena. It also uses the repeated games approach which has been used extensively in the commercial policy literature to explain aspects of international trade agreements. This paper shows the potential of linking these two traditions for modeling the geography of trade agreements.

6. Appendix

Proof of Proposition 3: $E(\sigma,s) \geq T(\sigma,s)$ can be rewritten as,

$$(1-\delta)G \geq \beta s\left\{\theta + 1 + \left[\theta\sigma \big/ \delta\left(\tfrac{1-\gamma}{\gamma-\delta} + \sigma\right)\right]\right\} \tag{A1}$$

the right-hand side of which is increasing in σ. Thus, if this condition holds at $\sigma = 1$, then it holds at any σ. By definition any $s \leq \bar{s}$ implies (A1) is satisfied at $\sigma = 1$. Hence, free trade can be sustained along any path to the free trade steady state. Since free trade is PE, it follows that free trade in every period is the unique PE equilibrium.

If $s \geq \hat{s}$, then (A1) cannot be satisfied for any $\sigma > 1-\gamma$. Recall that (A1) is the condition for countries to adopt free trade, given that they expect free trade in all future periods. However, if (A1) is violated, this expectation cannot be correct. Whatever alternative path is expected in the future, it cannot have as high a payoff as eternal free trade, and thus equilibrium enforcement must be lower than $E(\sigma_\tau, s)$. If free trade cannot be sustained by $E(\sigma_\tau, s)$, then it cannot be sustained by any lesser enforcement either. Hence, free trade cannot be sustained in any period, and the unique PE equilibrium must be $\hat{t}$ (or higher).

Finally, if $\bar{s} < s < \hat{s}$, there may be a cycle. Consider the simplest cycle: in every odd period countries adopt free trade; in every even period they adopt tariffs high enough to induce multi-plant production. For free trade to be an equilibrium in odd periods, it must be that, $s \leq s^{c(1)} \equiv G\delta^2 \big/ \beta(\theta\gamma^2 + \delta^2)$. Whether or not $\bar{s} < s^{c(1)} < \hat{s}$ depends on model parameters. Take for example, $\theta = 2$, $\delta = .5$ and $\gamma = .9$. Then $\hat{s}/\bar{s} = 1.335$ and $s^{c(1)}/\bar{s} = 1.175$, which implies there is a small range of transport costs consistent with a one-period cycling equilibrium. This Pareto dominates the high tariff steady state, though it may be dominated by a longer cycle if one exists.

Proof of Lemma: Suppose there exists an efficient equilibrium in which $t_{ij} > 0$ and exports from j to i are positive. Consider the effect of reducing t_{ij} to t'_{ij} $\quad (0, t_{ij})$. There are two possible cases:

(1) Suppose the reduction has no effect on the location of plants. Since $\tilde{\pi}$ and v are both decreasing in the tariff, the reduction in t_{ij} raises the consumer welfare in i and profits in j and causes no decrease in any component of any country's welfare. Thus, t'_{ij} represents a pareto improvement and it is also an equililbrium, as $\bar{V} - V$ does not rise and $\Pi - \bar{\Pi}$ does not fall for any country.

(2) Suppose the reduction in t_{ij} affects plant location. The firms that alter their plant location must be those headquartered in country $k \neq i, j$. Firms headquartered in i are entirely unaffected by a reduction in t_{ij}. By construction, firms headquartered in j export to country i from their home plants in the original equilibrium, a decision that is only reinforced by a reduction in t_{ij}. Thus, any alteration in plant location must be in the form of country k firms opening new plants in j so as to export from there to i. If such firms previously had no foreign plants, then the move raises profits in k and increases consumer welfare in i and j. Again t'_{ij} represents a pareto improvement and it is consistent with equililbrium. The only other possibility is that country k firms have plants in i in the orignial equilbirum but abandon them for plants in j, when t_{ij} is reduced. This would only occur if $t'_{ij} \leq t_{ji}$, but what is t_{ji}? There are

two possibilities: a) There are positive exports from i to j in the original equilibrium. If this is so, then t_{ji} must be zero, because any lowering of t_{ji} does not affect plant location and is thus pareto improving and consistent with equilibrium. b) There are no exports from i to j in the original equilibrium. If true, then there are also no exports from i to k. But this must be pareto dominated by an equilibrium in which $t_{ki} = \bar{t}$, $t_{ji} = 0$ and there are positive exports from i to j.

The consequence of a) and b) is that because the original equilibrium is pareto efficient, t_{ji} must be zero. It follows that for any $t_{ij} > 0$, there exists $t'_{ij} \quad (0, t_{ij})$ that pareto dominates t_{ij} and is consistent with equilibrium. This contracts our original supposition. QED

Proof of Proposition 5: Comparing (15) and (18), we have $W_A^{FT} > W_A^{AB}$ if and only if $F/B > s'(\theta+1) - s(\theta/2)$. But (17) guarantees that $\frac{F}{\beta} > s'\left(\frac{\theta}{\delta}+1\right) > s'(\theta+1) - s\left(\frac{\theta}{2}\right)$, and so, $W_A^{FT} > W_A^{AB}$ whenever (17) is satisfied. Thus, we can rank the payoffs to country A, as follows:

$$W_A^{AH} > W_A^{FT} > W_A^{AB} > W_A^{AC} = W_A^{CH} \tag{A2}$$

As $W_A^{FT} = W_B^{FT}$ and $W_A^{AB} = W_B^{AB}$, we can partially rank the payoffs for country B:

$$W_B^{FT} > W_B^{AB} = W_B^{AH} > W_A^{CH} \tag{A3}$$

From A2 and A3, we see that FT and AH cannot be Pareto dominated. This establishes proposition 5a. Also, FT (and only FT) is better than AB for both A and B. Thus, AB is Pareto dominated if and only if $W_C^{FT} > W_C^{AB}$. Subtracting (19) from (16), a little algebra produces the condition: $s \le 2(\theta+1)s' - (F/\beta)$. This establishes 5b. QED

Proof of Proposition 6: For all paramters W_C^{AC} is the lowest payoff (other than the punishment equilibrium) for country C. Thus, considering A2, AC must be Pareto dominated unless $W_B^{AC} > W_B^{FT}$. Subtracting (21) from (15), a little algebra produces the condition: $s \ge \frac{(F/\beta) - \theta s'}{1+\theta}$. This establishes 6a.

A2, A3, and the fact that $W_B^{AC} > W_B^{CH}$ for all paramters imply that CH is the worst outcome for A and B. Thus, CH is Pareto dominated unless it is the best outcome for C. For all parameters it is the case that $W_C^{CH} > W_C^{FT}$ and $W_C^{AB} > W_C^{AH} > W_C^{AC}$. Thus the CH is best for C if and only if $W_C^{CH} > W_C^{AB}$, which if produces the condition: $s \ge (\theta+2)s' - (F/\beta)$. This establishes 6b.

Finally, we note that condition (23) depends on the discount factor. In particular if $\delta \le 3/4$, then

$$\frac{F}{\beta} \ge s'\left(\frac{3}{2}\frac{\theta}{\delta}+1\right) \ge s'(2\theta+1) \tag{A3}$$

Recalling that $\theta > 1$, and $s < s'$, A3 ensures the condition in 6b is satisfied and that in 6a is violated. Thus, CH is a PE equilibrium and AC is not. QED

7. References

Abreu, D., 1988, On the theory of infinitely repeated games with discounting, Econometrica 56(2), 383-396.

Baldwin, R. E. and A. J. Venables, 1995, Regional economic integration, in: G. Grossman and K. Rogoff, eds., Handbook of international economics, Vol. III (North-Holland, Amsterdam).

Bagwell, K. and R. Staiger, 1990, A theory of managed trade, American Economic Review 80, 779-795.

Bagwell, K. and R. Staiger, 1997a, Multilateral tariff cooperation during the formation of customs unions, Journal of International Economics 42, 91-124.

Bagwell, K. and R. Staiger, 1997a, Multilateral tariff cooperation during the formation of free trade areas, International Economic Review 38(2), 291-320.

Bagwell, K. and R. Staiger, 1999, An economic theory of GATT, American Economic Review, 89(1), 215-48.

Blonigen, B. and R. Feenstra, 1997, Protectionist threats and foreign direct investment, in: R. Feenstra, ed., The effects of US trade protection and promotion policies (University of Chicago Press, Chicago, IL) 55-80.

Brainard, S. L., 1993, A simple theory of multinational corporations and trade with a trade-off between proximity and concentration, NBER working paper 4269.

Brainard, S. L., 1997, An empirical assessment of the proximity-concentration trade-off between multinational sales and trade, American Economic Review 87(4), 521-544.

Brander, J. S. and B. J. Spencer, 1987, Foreign direct investment with unemployment and endogenous taxes and tariffs, Journal of International Economics 22, 257-79.

Bond, E., 1999, Multilateralism vs. regionalism: tariff cooperation and interregional transport costs, in: S. Lahiri, Regionalism and multilateralism (Routledge Press, New York).

Bond, E. and C. Syropoulos, 1996, Trading blocs and the sustainability of interregional cooperation, in: M. Canzoneri, W. Ethier and V. Grilli, eds., The new transatlantic economy (Cambridge University Press, New York) 118-141.

Brecher, R. A. and C. Diaz Alejandro, 1997, Tariffs, foreign capital and immiserizing growth, Journal of International Economics 7, 317-22.

Caves, R., 1996, Multinational enterprise and economic analysis, 2nd ed. (Cambridge University Press, New York).

Etheir, W. J., 1999, Regional regionalism, mimeo, University of Pennsylvania.

Frankel, J., E. Stein and S. Wei, 1997, Trading blocs and the Americas: the natural, the unnatural, and the super-natural, Journal of Development Economics 47(1), 61-95.

Helpman, E. and P. Krugman, 1989, Trade policy and market structure (MIT Press, Cambridge, MA).

Horst, T., 1971, The theory of the multinational firm: optimal behavior under different tariff and tax rules, Journal of Political Economy 79, 1059-72.

Horstmann, I. and J. R. Markusen, 1992, Endogenous market structures in international trade, Journal of International Economics 32, 109-129.

Hummels, D., 1998, Transportation costs and international integration in recent history, mimeo, University of Chicago.

Krugman, P. R., 1983, The ‘new theories’ of international trade and the multinational enterprise, in: C. Kindleberger and D. Audretsch, eds., The multinational corporation in the 1980s (MIT Press, Cambridge, MA).

Krugman, P. R., 1991a, The move toward free trade zones, in: Policy implications of trade and currency zones : a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming,August 22-24, 1991(Federal Reserve Bank of Kansas City, Kansas City, MO).

Krugman, P. R., 1991b, Increasing returns and economic geography, Journal of Political Economy 99, 183-199.

Krugman, P. R., 1992, Geography and Trade (MIT Press, Cambridge, MA).

Ludema, R. D., 1999, Why are preferential trade agreements regional? increasing returns, multinationals and the geography of free trade agreements,” Georgetown University Working Paper 99-04.

Maggi, G., 1999, The role of multilateral institutions in international trade cooperation, American Economic Review 89(1), 190-214.

Markusen, J. R., 1995, The boundaries of multinational enterprises and the theory of international trade, Journal of Economic Perspectives 9(2), 169-189.

Markusen, J. R. and A. J. Venables, 1998, Multinational firms and the new trade theory, Journal of International Economics, 46, 183-203.

Motta, M., 1992, Multinational firms and the tariff-jumping argument: a game-theoretic analysis with some unconventional conclusions, European Economic Review 36, 1557-1571.

Smith, A., 1987, Strategic investment, multinational corporations and trade policy, European Economic Review 31, 89-96.

Srinivasan, T. N., 1983, International factor movements, commodity trade and commercial policy in a specific factors model, Journal of International Economics 14, 289-312.

Staiger, R., 1995, International rules and institutions for cooperative trade policy in: G. Grossman and K. Rogoff, eds., Handbook of international economics, Vol. III (North-Holland, Amsterdam).

Venables, A. J., 1985, Trade and trade policy with imperfect competition: the case of identical products and free entry, Journal of International Economics 19, 1-20.

Venables, A. J., 1987, Trade and trade policy with differentiated products: a chamberlinian-ricardian model, Economic Journal 97, 700-717.

Viner, J., 1950, The customs union issue (Carnegie Endowment, New York).

Wonnacott, R. J., 1990, U.S. hub-and-spoke bilaterals and the multilateral trading system, C. R. Howe Institute Commentary, No. 23.