The Economic Journal, **108** (*July*), 1162–1182. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

WILL PREFERENTIAL AGREEMENTS UNDERMINE THE MULTILATERAL TRADING SYSTEM?*

Kyle Bagwell and Robert W. Staiger

Is bilateralism bad? This question was the title of Paul Krugman's provocative paper on regionalism (Krugman, 1991*a*), and the pursuit of an answer has spawned the *regionalism versus multilateralism* debate. The answer to this question turns largely on how one answers the related question posed by Jagdish Bhagwati (Bhagwati, 1991): will regional agreements – or more accurately, *preferential* agreements – undermine the multilateral trading system? On one side of the debate are those who argue that preferential agreements can complement existing multilateral efforts to foster greater economic integration among countries, and should therefore be encouraged. On the other side are those who see such agreements as a threat to the multilateral system.

But what is 'the multilateral system' that may or may not be threatened by preferential agreements? Much of the literature treats it as a 'black box' synonymous with the goal of 'multilateral free trade', and proceeds to ask whether preferential agreements contribute to or interfere with the attainment of this goal. Introductory references are typically made to the GATT/WTO, but the analysis is often carried out with little or no reference to the structure of this multilateral institution. In this paper we argue that understanding GATT's structure is vital to the debate over regionalism versus multilateralism.

There are three components that together comprise the cornerstones of the GATT system: the principles of *reciprocity* and *non-discrimination*, which are regularly identified as the 'pillars' of the GATT architecture, and the *enforcement* mechanisms, which Dam (1970, p. 81) calls the 'heart' of the GATT system. Below we describe how preferential agreements may be expected to interact with the multilateral system in light of each of its three principle components.

We approach the question posed at the outset in two parts. We first ask, will preferential agreements undermine a multilateral trading system that is built on the pillars of reciprocity and non-discrimination? A remaining question is then, how do preferential agreements affect the enforcement provisions of the GATT? To lay the foundation for an answer to the first question, we describe a framework within which the pillars of reciprocity and non-discrimination can themselves be interpreted and understood. From this perspective, we then

* This paper was completed while Staiger was a Fellow at the Center for Advanced Study in the Behavioral Sciences. Staiger is also grateful for financial support provided by The National Science Foundation Grant #SBR-9022192.

offer support for the view that preferential agreements pose a threat to the multilateral system. An answer to the second question requires an understanding of how GATT agreements are enforced. Observing that these agreements must be *self-enforcing*, we describe circumstances under which preferential agreements can either enhance or detract from the performance of the GATT system through their impacts on enforcement at the multilateral level.

In order to establish these answers, we must first articulate a theory of the multilateral trading system within which the workings of reciprocity, non-discrimination and enforcement mechanisms can be understood. Any attempt to construct such a theory immediately confronts a most-basic question: Why have governments found reciprocal trade agreements to be appealing? One view is that these agreements provide governments with an escape from a terms-of-trade driven Prisoners' Dilemma. This view has a long history, originating with Torrens (1844), Mill (1844), Scitovsky (1942) and Johnson (1953–4), who posed their arguments in the context of national-income maximising governments that set 'optimal tariffs' to exploit monopoly power in world markets.

While the terms-of-trade view is logically correct, many economists are sceptical of its practical relevance: a common position is that these arguments become secondary in explaining the appeal of reciprocal trade agreements once more realistic government objectives that include political concerns are introduced. A popular view is then that governments are attracted to such agreements as a means to achieve political objectives.

To evaluate these two views, we therefore begin in the next section by constructing a general model in which governments are motivated by both political and terms-of-trade considerations. Our framework is sufficiently general to include all the major political economy models of trade policy formulation as special cases. Within this framework, we show that more general government objectives do not change the view that trade agreements provide an escape from a terms-of-trade driven Prisoners' Dilemma, and that this is *all* that trade agreements do. Political forces shape the trade objectives that governments seek to achieve, but political considerations play no role in explaining why governments seek reciprocal trade agreements. Rather, it is the terms-of-trade externality that creates an inefficiency when policies are set unilaterally and that therefore explains why governments need a reciprocal trade agreement in order to accomplish their objectives.[1]

As it turns out, this observation both clarifies the purpose of trade agreements in the leading models of trade policy and, as we will demonstrate, reveals a simple logic to GATT's principles of reciprocity and non-discrimina-

[1] Political considerations might provide a separate motivation for reciprocal trade agreements if governments seek such agreements to gain commitment relative to their private sectors, a possibility explored by Maggi and Rodriguez (forthcoming) and Staiger (1995). It has not yet been demonstrated, however, whether this hypothesis could account for the principles of reciprocity and non-discrimination that form the pillars of GATT.

tion. Furthermore, with this theory of the multilateral trading system at hand, we are then able to offer our answers concerning the consequences of preferential trading agreements for the functioning of the multilateral trading system.

1. The Purpose of Reciprocal Trade Agreements

We develop here a general model of the trade policy choices of politically motivated governments. Our initial goal is to understand the appeal of reciprocal trade agreements for such governments. We present a full treatment of the issues contained in this and the next section in Bagwell and Staiger (1997a).

1.1. *The Economic Environment*

We first describe the economic environment, which is a standard two-good two-country general equilibrium trade model. A home (no *) and foreign (*) country trade two goods, x and y, which are normal goods in consumption and produced in competitive markets under conditions of increasing opportunity costs. With x (y) the natural import good of the home (foreign) country, define $p = p_x/p_y$ to be the local relative price facing home producers and consumers, and similarly define $p^* = p_x^*/p_y^*$ to be the local relative price facing foreign producers and consumers. Letting τ denote one plus the (non-prohibitive) *ad valorem* import tariff imposed by the home country, and defining τ^* analogously for the foreign country, it then follows that $p = \tau p^W \equiv p(\tau, p^W)$ and $p^* = p^W/\tau^* \equiv p^*(\tau^*, p^W)$, where $p^W \equiv p_x^*/p_y$ is the 'world' (i.e., untaxed) relative price. The home country (foreign country) terms of trade is then given by $1/p^W(p^W)$.

Within each country, local relative prices determine competitive production decisions, which we represent by the domestic and foreign production functions $Q_i = Q_i(p)$ and $Q_i^* = Q_i^*(p^*)$ for $i = \{x, y\}$, respectively. National consumption is determined by the local relative price, which defines the tradeoff faced by consumers and implies the level and distribution of factor income in the economy, and by tariff revenue, which is distributed to consumers lump-sum. In the usual way, tariff revenue can be expressed as a function of local and world prices, so that we may represent national consumption in each country as $C_i(p, p^W)$ and $C_i^*(p^*, p^W)$. Finally, home country imports of x, denoted by $M_x(p(\tau, p^W), p^W)$, are given by the difference between home country consumption and production of x, while home country exports of y, denoted by $E_y(p(\tau, p^W), p^W)$, amount to the difference between home country production and consumption of y. Foreign country imports of y, $M_y^*(p^*(\tau^*, p^W), p^W)$, and exports of x, $E_x^*(p^*(\tau^*, p^W), p^W)$, are similarly defined.

We may now express the trade balance and equilibrium conditions. For any world price, home and foreign budget constraints imply that

$$p^W M_x(p(\tau, p^W), p^W) = E_y(p(\tau, p^W), p^W) \text{ and} \quad (1a)$$

$$M^*_y(p^*(\tau^*, p^W), p^W) = p^W E^*_x(p^*(\tau^*, p^W), p^W). \quad (1b)$$

The equilibrium world price, $\tilde{p}^W(\tau, \tau^*)$, is then required to clear the market for good y:

$$E_y(p(\tau, \tilde{p}^W), \tilde{p}^W) = M^*_y(p^*(\tau^*, \tilde{p}^W), \tilde{p}^W), \quad (2)$$

with market clearing for good x then implied by (1) and (2). Hence, given any pair of non-prohibitive tariffs, the equilibrium world price will be determined through (2), and the equilibrium world price and the given tariffs then determine in turn the local prices and thereby production, consumption, import, export and tariff revenue levels. Throughout we assume that the Metzler and Lerner paradoxes are ruled out, so that $dp/d\tau > 0 > dp^*/d\tau^*$ and $\partial \tilde{p}^W/\partial\tau < 0 < \partial \tilde{p}^W/\partial\tau^*$.

1.2. *Government Objectives*

We adopt a representation of government objectives which is general enough to include as special cases national income maximisation as well as each of the major modelling approaches to the political economy of trade policy. The key observation is that politically motivated governments are sensitive to the distributional as well as the efficiency properties of the local price implications of their trade policy choices. We thus represent the objectives of home and foreign governments by the general functions $W(p(\tau, \tilde{p}^W), \tilde{p}^W)$ and $W^*(p^*(\tau^*, \tilde{p}^W), \tilde{p}^W)$, respectively. The only structure we place on these functions is that, holding local prices fixed, each government is assumed to achieve higher welfare when its terms-of-trade improve:

$$\partial W(p, \tilde{p}^W)/\partial \tilde{p}^W < 0 \text{ and } \partial W^*(p^*, \tilde{p}^W)/\partial \tilde{p}^W > 0. \quad (3)$$

For fixed local prices, an improvement in a country's terms of trade amounts to an income transfer from its trading partner (brought about by an increase in the country's tariff and a corresponding decrease in the tariff of its trading partner). Throughout, we also assume that the second-order conditions associated with the maximisation problems developed below are globally satisfied.

1.3. *Unilateral Trade Policies*

Consider first the unilateral trade policies that governments would choose in the absence of a trade agreement. Supposing that each government would set its trade policy to maximise its objective function taking as given the tariff choices of its trading partner, the associated home and foreign reaction functions are defined implicitly by:

$$\text{Home:} \quad W_p + \lambda W_{p^w} = 0, \quad (4a)$$

$$\text{Foreign:} \quad W^*_{p^*} + \lambda^* W^*_{p^w} = 0, \quad (4b)$$

where subscripts denote partial derivatives, and where $\lambda \equiv (\partial \tilde{p}^W/\partial\tau)/(dp/d\tau) < 0$ and $\lambda^* \equiv (\partial \tilde{p}^W/\partial\tau^*)/(dp^*/d\tau^*) < 0$. As (4) illustrates, the best-response tariff of each government is determined by the combined impact on welfare of the induced local and world price movements.

The best-response conditions can be further interpreted with the aid of Fig. 1. An initial tariff pair is given by point $A \equiv (\tau, \tau^*)$. Associated with the point A is an iso-local-price locus, denoted as $p(A) \rightarrow p(A)$, and an iso-world-price locus, labelled $p^W(A) \rightarrow p^W(A)$.[2] If the home government were to raise its tariff unilaterally to τ^1, then a new tariff pair would be induced, represented by

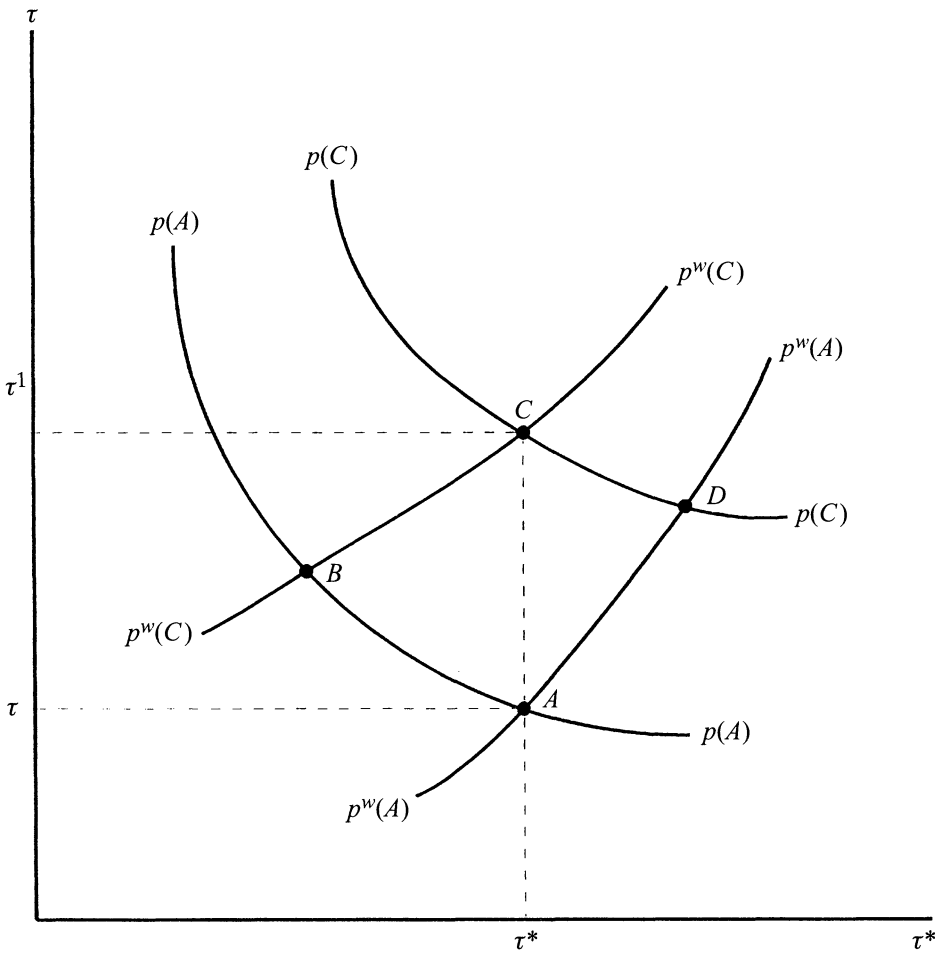

Fig. 1.

[2] Excluding the Metzler and Lerner paradoxes, the iso-local-price locus is negatively sloped while the iso-world-price locus has positive slope.

the point $C \equiv (\tau^1, \tau^*)$ in Fig. 1, and this tariff pair rests on a new iso-local-price locus $p(C) \rightarrow p(C)$ and also a new iso-world-price locus $p^W(C) \rightarrow p^W(C)$. Hence, with a unilateral increase in its tariff, the home government induces a local price that is higher and a world price that is lower as compared to the prices associated with the original point A.

In analogy with $(4a)$, Fig. 1 can be used to disentangle the overall movement from A to C induced by a unilateral tariff increase by the domestic government into separate movements in the local and world prices, respectively. The movement from A to B isolates the induced reduction in the world price, and its welfare significance for the domestic government is captured in $(4a)$ by the term W_{p^W}. Similarly, the movement from B to C isolates the local price change, and its welfare significance for the domestic government is captured in $(4a)$ by the term W_p. The welfare implications of the local-price change implied by the movement from B to C reflect a weighing of the costs of the associated domestic distortions in production and consumption against any domestic distributional benefits. The welfare implications of the change in the world-price implied by the movement from A to B, in contrast, reflect the benefits accruing to the domestic government as it shifts the costs of its policy onto the foreign government through a terms-of-trade improvement. It follows that, if the home government seeks to implement a local price corresponding with the iso-local-price locus $p(C) \rightarrow p(C)$, then a unilateral increase in the domestic import tariff will shift a portion of the costs of achieving this policy goal onto the foreign government. A similar interpretation holds for $(4b)$.

A tariff pair that simultaneously satisfies $(4a)$ and $(4b)$ constitutes a Nash equilibrium of the game that arises when both governments set tariffs unilaterally. We will assume that such a tariff pair, which we take to be unique, corresponds to the tariff choices that governments would make in the absence of a trade agreement, and we will denote this tariff pair by (τ^N, τ^{*N}). The next task is to determine whether these tariff choices are inefficient for the governments making them and, if they are, to identify the source of the inefficiency that a trade agreement can correct.

1.4 *Why Unilateral Trade Policy Choices are Inefficient*

To determine whether the tariff choices made by governments in the absence of a trade agreement are efficient for the governments given their objectives, we need to characterise the set of tariff pairs that lie on the efficiency frontier and ask whether (τ^N, τ^{*N}) is an element of this set. The efficiency frontier can be given the general representation $(d\tau/d\tau^*)|_{dW=0} = (d\tau/d\tau^*)|_{dW^*=0}$, but it can also be represented more concretely as the set of tariffs that satisfy:

$$(1 - AW_p)(1 - A^* W^*_{p^*}) = 1, \tag{5}$$

where $A \equiv (1 - \tau\lambda)/(W_p + \lambda W_{p^w})$ and $A^* \equiv (1 - \lambda^*/\tau^*)/(W^*_{p^*} + \lambda W^*_{p^w})$,

with $A \neq 0$ and $A^* \neq 0$ under the further assumption that the partial derivatives of the welfare functions are always finite.[3]

It is now immediate from a comparison of (4) and (5) that Nash tariffs are inefficient. This is not surprising since, as we have described above, when governments set their trade policies unilaterally they are motivated to shift costs onto one another through the world-price changes that their tariffs imply. It is perhaps also not surprising that these cost-shifting motives will lead governments to adopt trade policies that are unambiguously too restrictive relative to efficient choices given their objectives: it can be shown that both governments can achieve welfare gains relative to the Nash equilibrium only if each agrees to lower its tariff below the Nash level.[4]

What is perhaps more surprising is that this terms-of-trade externality is the *only* inefficiency that a trade agreement can remedy, despite the fact that we have allowed governments to have political motivations of a quite general nature. To establish this, we consider a hypothetical world in which governments are not motivated by the terms-of-trade consequences of their trade policies. To this end, we define *politically optimal tariffs* as any tariff pair (τ^{PO}, τ^{*PO}) satisfying the following two conditions:

$$\text{Home:} \qquad W_p = 0 \tag{6a}$$

$$\text{Foreign:} \qquad W^*_{p^*} = 0. \tag{6b}$$

We assume that a unique set of politically optimal tariffs exists and that the associated second-order conditions are globally satisfied.

When governments set politically optimal tariffs, it is as if they ignore any welfare gains attributable to changes in the world price that would be induced by their tariff choices. If both governments sought to maximise national income, politically optimal tariffs would correspond to reciprocal free trade. More generally, however, government objectives may reflect political considerations as well, and in this case there is no presumption that politically optimal tariffs will correspond to free trade.

It can now be seen that politically optimal tariffs are efficient, as tariffs that satisfy (6) will lie along the efficiency locus defined by (5). As a consequence, once the terms-of-trade motivations are eliminated from the trade-policy choices of governments, there is no further scope for Pareto improving changes in trade policy. To gain intuition for this finding, suppose that tariffs

[3] If governments maximise national income, the efficiency locus reduces to the form $\tau = 1/\tau^*$, as Mayer (1981) demonstrates. In this case, tariffs are adjusted along the efficiency locus so as to maintain equality in relative local prices between the domestic and foreign countries, with different tariff pairs along the efficiency locus reflecting different world prices and therefore different distributions of income across trading partners. In the more general representation of government preferences considered here, the efficiency locus still determines a relationship between domestic and foreign tariffs, but it need no longer be the case that this relationship equates relative local prices across trading partners.

[4] A reduction in tarrifs from the Nash level, however, is not sufficient to guarantee mutual welfare gains. For example, as Johnson (1953–4) and Kennan and Riezman (1988) demonstrate, a large country may be worse off under reciprocal free trade than in the Nash equilibrium if countries are asymmetric.

have been set at their politically optimal levels, so that the terms-of-trade motivations have been removed from trade-policy choices and each government has set its trade policy so as to achieve its preferred local prices. Consider now a small increase in the tariff of the domestic country. This change has three effects. First, it induces a small increase in the local price in the domestic economy, but this has no first-order effect on domestic government welfare as domestic local prices were already at their preferred level. Second, the domestic tariff increase induces a small reduction in the local price of the foreign country, but this has no first-order effect on the welfare of the foreign government since it too has already achieved its preferred local prices. Finally, the domestic tariff increase causes the world price to fall, but this cannot generate an efficiency gain as it represents a pure international transfer of tariff revenue. Hence, once the terms-of-trade motivations have been removed from trade-policy choices, there are no further Pareto gains for governments to achieve.

Fig. 1 illustrates the essential inefficiency that prevents governments from reaching the efficiency frontier with unilateral trade policy decisions. We suppose again that, beginning at point A, the home government considers a unilateral tariff increase to achieve the local price associated with point C. If the government allows the terms-of-trade consequences (i.e., the movement from D to C) of its tariff selection to influence its decision concerning whether to proceed with this tariff increase, then it will recognise that some of the costs of achieving the higher local price at C can be shifted onto its trading partner as a result of the reduced world price, and the tariff increase will look especially attractive. As a consequence, Nash tariffs are always inefficient, leading to tariffs (trade volumes) that are too high (low). Alternatively, if the government were not permitted to let the terms-of-trade consequences of its tariff selection influence its decision of whether or not to proceed with the tariff increase, then it would prefer choosing the higher tariff to induce point C if and only if it also prefers point D to point A. In this case, the potential appeal of point C to the home government is independent of any cost-shifting benefits that may arise with the consequent change in the world price, and so it has the 'right' incentives when deciding whether to proceed with a tariff increase.[5] If each government were to choose tariffs in this way, then a resulting set of consistent tariffs is politically optimal and efficient.

The findings described here can be summarised with the aid of Fig. 2, which depicts the locus of efficient tariff pairs by the curve $E \rightarrow E$, the politically optimal tariffs by the point on the efficiency locus labeled PO, and the Nash tariffs by the point labelled N positioned to the northeast of the efficiency

[5] The movement from point A to point D in Fig. 1 induces no externality through the terms-of trade but it does alter the foreign local price. If the foreign government also selects tariffs that are politcally optimal, however, then a small change in the foreign local price will have no first-order effect on foreign government welfare.

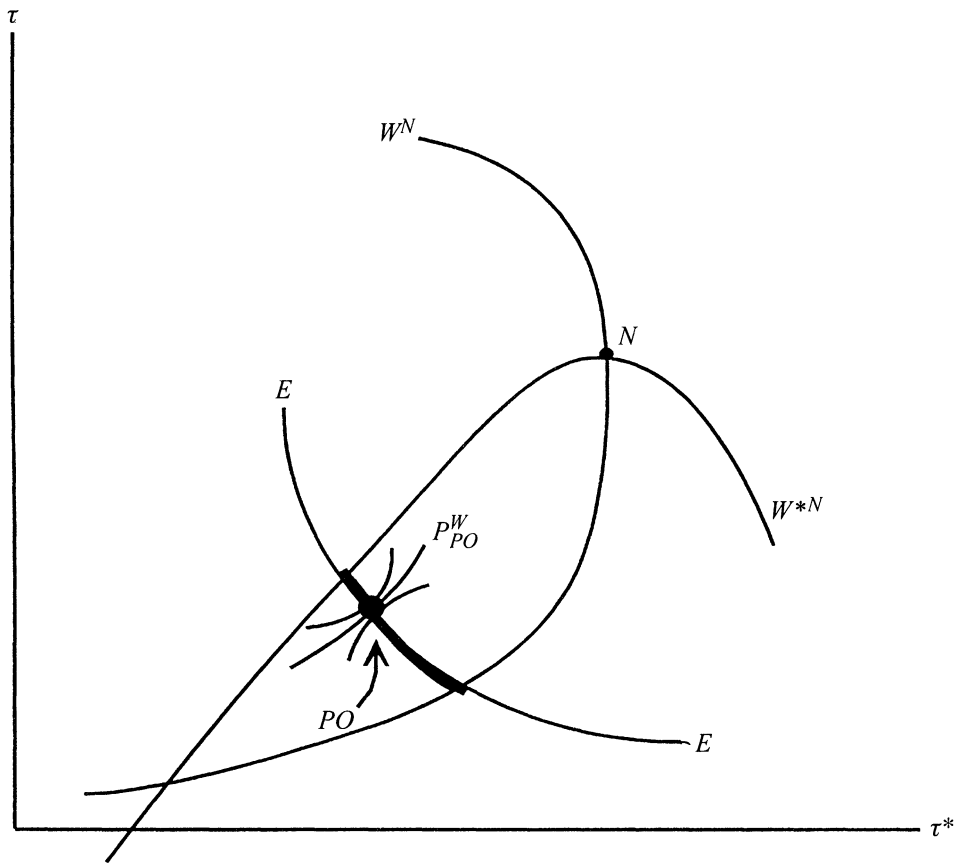

Fig. 2.

locus.[6] Notice that the iso-welfare curves of the two governments are tangent at every point along the efficiency locus, including the political optimum. The novel feature of the politically optimal tariffs is that the iso-welfare curves at these tariffs are also tangent to the iso-world-price locus (the locus labelled p^W_{PO}). The bold portion of the efficiency locus corresponds to the contract curve.

Fig. 2 clarifies the central task that governments face when they design a trade agreement. Without cooperation, governments would set trade policy unilaterally, leading to the Nash outcome *N*. A trade agreement can then offer governments a means to cooperate and move from the inefficient Nash point to some alternative tariff pair that rests on the contract curve. Among the points on the contract curve, the politically optimal tariffs are focal, as they

[6] We draw this picture under the assumptions that a unique Nash equilibrium exists, a unique political optimum exists, and that the political optimum lies on the contract curve (i.e., it is on the efficiency locus and yields greater-than-Nash welfare for each government). The last assumption is new to our discussion, and it will be satisfied if countries are sufficiently symmetric.

remedy in a direct fashion the terms-of-trade inefficiency that keeps governments from reaching the contract curve with their unilateral choices. As Fig. 2 makes clear, when governments have political objectives the efficiency locus need not pass through the point of reciprocal free trade. But while political concerns will effect government preferences over tariffs (e.g., the location of the efficiency locus), it is the terms-of-trade externality that creates a 'problem' with unilateral tariff choices which an appropriately designed trade agreement can 'solve'.[7]

2. Reciprocity, Non-discrimination and Preferential Agreements

With the basic framework now described, we are prepared to consider the principles of reciprocity and non-discrimination, and how preferential agreements will affect a multilateral trading system built on these two pillars.

2.1. *The Meaning of Reciprocity*

Within GATT, the term 'reciprocity' refers broadly to the ideal of mutual changes in trade policy that bring about equal changes in import volumes across trading partners.[8] Using the model presented above, we may define reciprocity more formally as follows: a set of tariff changes $\Delta\tau \equiv (\tau^1 - \tau^0)$ and $\Delta\tau^* \equiv (\tau^{*1} - \tau^{*0})$ conforms to *the principle of reciprocity* provided that

$$\tilde{p}^{W0}[M_x(p(\tau^1, \tilde{p}^{W1}), \tilde{p}^{W1}) - M_x(p(\tau^0, \tilde{p}^{W0}), \tilde{p}^{W0})] = [M_y^*(p^*(\tau^{*1}, \tilde{p}^{W1}), \tilde{p}^{W1}) - M_y^*(p^*(\tau^{*0}, \tilde{p}^{W0}), \tilde{p}^{W0})]$$

where $\tilde{p}^{W0} \equiv \tilde{p}^W(\tau^0, \tau^{*0})$, $\tilde{p}^{W1} \equiv \tilde{p}^{W1}(\tau^1, \tau^{*1})$, and where we have measured changes in import volumes at existing world prices. Using the trade balance condition (1) and the equilibrium condition (2), this expression may be rewritten as

$$[\tilde{p}^{W1} - \tilde{p}^{W0}]M_x(p(\tau^1, \tilde{p}^{W1}), \tilde{p}^{W1}) = 0. \tag{7}$$

[7] The degree to which countries are able to affect their terms of trade significantly is an issue of some debate. We note the following points of support. First, at the level of theory, even ostensibly small countries have some power over their terms of trade, if the industry is monopolistically competitive (Gros, 1987). Second, our theory does not require that all countries have the ability to alter world prices, but it does imply that only those countries that can alter world prices will be actively involved in reciprocal tariff negotiations, an implication that is consistent with the 'principal supplier' rules of GATT negotiations (see Bagwell and Staiger, 1996). Third, with regard to empirical evidence, a large literature documents imperfect 'pass-through' in the face of exchange rate shocks. If symmetric empirical patterns arise when the cost increase is associated with a tariff increase, then the finding of imperfect pass-through would offer evidence of a reduction in the world price, i.e., a terms-of-trade externality. Feenstra (1989) provides empirical support for the symmetric pass-through hypothesis. Finally, we note that empirical studies of trade policy confirm that the potential world-price implications of alternative trade policy choices can have important effects on the national desirability of intervention (see, for example, the discussion in Feenstra, 1995, p. 1579).

[8] The meaning of reciprocity in GATT, and the various ways in which reciprocity has been implemented in practice, is discussed in Dam (1970, pp. 58–61 and pp. 87–91) and WTO (1995a, p. 949).

Hence, as (7) makes clear, mutual tariff changes that conform to reciprocity leave world prices unchanged. The potential significance of this property can be appreciated when viewed from the perspective of the finding, reported above, that a government's tariff choice will be inefficient if and only if it is motivated by the *change* in the world price that its tariff choice implies.

2.2. *The Practice of Reciprocity in GATT*

Having now defined the general meaning of the principle of reciprocity, we turn to the application of this principle within GATT practice. To understand the importance of the principle of reciprocity in GATT, it is useful to distinguish between two broad circumstances in which reciprocity applies. A first circumstance is when governments seek greater access to the markets of their trading partners, and engage in a 'round' of negotiations under GATT's Article XXVIII bis. In the context of these negotiations, opening one's own market is deemed a 'concession'. In this circumstance, the principle of reciprocity reflects the 'balance of concessions' that governments seek through a negotiated agreement. This practice is described by Dam (1970, p. 59), who explains that, under the language of Article XXVIII bis, negotiations are voluntary and are to be conducted in a 'reciprocal and mutually advantageous basis'. Dam (1970, p. 59) explains further that:

> 'This permissive approach to the content of tariff agreements is often referred to under the heading of *reciprocity*. From the legal principle that a country need make concessions only when other contracting parties offer reciprocal concessions considered to be "mutually advantageous" has been derived the informal principle that exchanges of concessions must entail reciprocity'.

The emphasis that governments place on reciprocity in this sense stands in contrast to standard economic logic, which holds that optimal unilateral policy for a country is free trade. From this perspective, it is perplexing that a government would require a 'concession' from its trading partner in order to do what is in any event best for its country. Appealing to this apparent violation of economic logic, it is tempting to interpret the observation that governments seek reciprocity in negotiated agreements as direct evidence that government negotiators adopt a mercantilist perspective that is incompatible with basic economic reasoning and that therefore derives from underlying political forces.[9]

Here we simply note that the mercantilist logic that drives actual trade negotiations admits a simple economic interpretation within the framework developed in the preceding section. In particular, the nature of the terms-of-

[9] Many have expressed this view. Krugman (1991*b*, 1997) provides an especially clear articulation of this position.

trade externality described above ensures that, in the absence of a trade agreement (i.e., at the Nash point), each government would prefer to reduce its tariff and induce lower import-competing prices and greater imports *if the increased trade volume could be obtained without a deterioration in the terms of trade* (see (4)).[10] Acting unilaterally, a government cannot achieve this. But by balancing one country's 'concessions' against another's, this is precisely what liberalisation conforming to the principle of reciprocity will deliver.

While reciprocity in this circumstance reflects the broad manner in which governments appear to approach rounds of trade negotiations under Article XXVIII bis, there is in fact no requirement in GATT that negotiations proceed in this manner. There is, however, a second broad circumstance in which the principle of reciprocity applies in GATT practice, and in this case GATT rules do require reciprocity. Here we refer to the circumstance in which a government wishes to reduce foreign access to its markets below a previously negotiated level. An important mechanism through which GATT provides for this possibility is contained in Article XXVIII.[11] Under this article, a country may at any time propose to modify or withdraw a concession agreed upon in a previous round of negotiation. In this circumstance, if the country and its trading partner cannot agree on a renegotiated tariff structure, then the country is free to carry out the proposed changes to its tariffs, and the notion of reciprocity is then invoked to moderate the allowable response of the country's trading partner, who is permitted to withdraw *substantially equivalent concessions* of its own.

This suggests that GATT negotiations may be understood as a multi-stage game, in which governments first agree to an initial set of tariffs, and then each government considers whether to propose a more restrictive tariff knowing that, under GATT's reciprocity rules, its trading partner can do no more than respond with reciprocal withdrawals of concessions which preserve the world price implied by the original agreement. Viewed from this perspective, it is clear that when governments evaluate the desirability of a proposed initial agreement, they must take account of any future incentives to renegotiate the agreement under GATT's reciprocity rules. But it should now also be clear that an initial agreement will be renegotiated if either government decides that the protection from imports provided under the agreement is inadequate at the existing world price. The implications of this observation can be illustrated with the help of a figure.

Consider then Fig. 3, which depicts three possible tariff pairs that might represent an initial agreement. We represent these tariff pairs by the points *A*, *B*, and *PO*. Each tariff pair is efficient (all three points lie on the efficiency locus), but the political optimum point PO is distinguished from the others by

[10] This follows from (3) and (4*a*) which indicate that $W_p < 0$ along the domestic government's reaction curve. A similar observation holds for the foreign government.

[11] For further discussion of Article XXVIII and its importance in GATT practice, see Dam (1970, pp. 79–99), Jackson (1989, p. 119) and Enders (1997). Provisions for temporary suspension of GATT obligations are provided in Article XIX, where similar reciprocity rules apply.

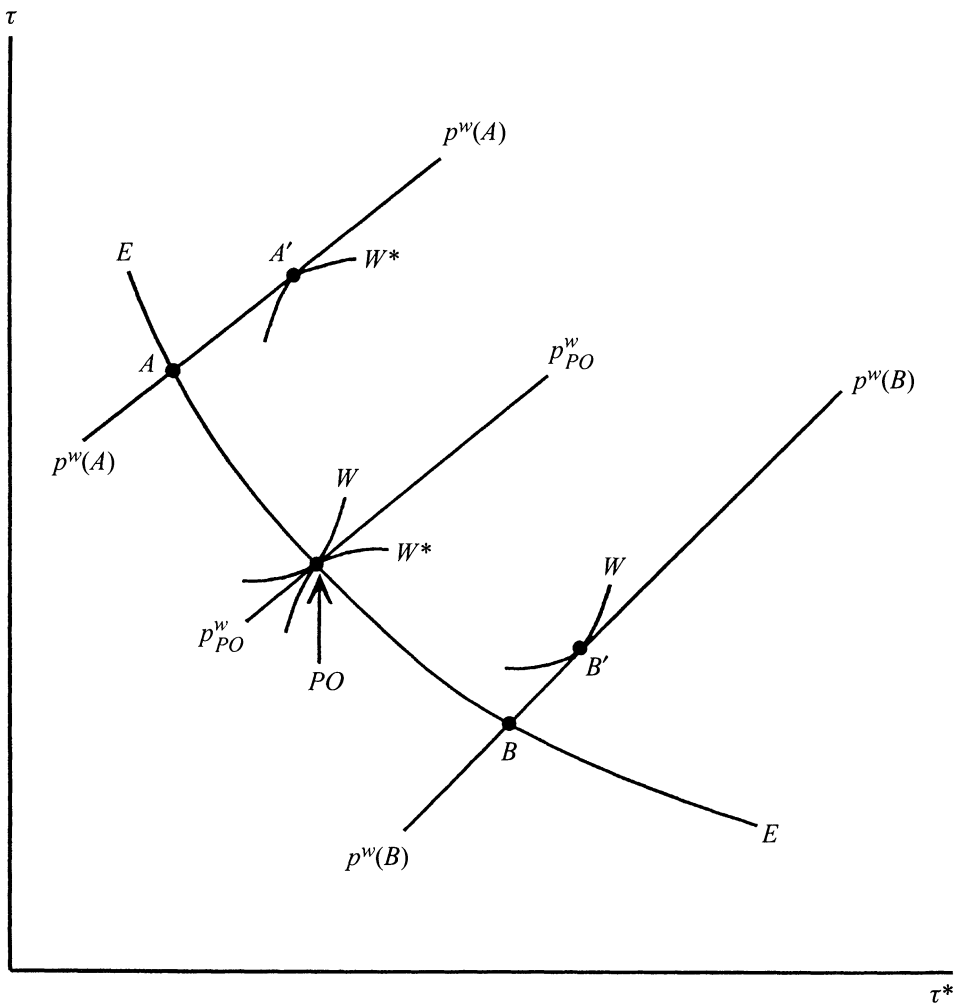

Fig. 3.

the fact that the iso-world-price locus running through *PO* is tangent to the iso-welfare contours of each government at this point.

Consider first an initial agreement that corresponds to point *A*. Observe that in this case the foreign government would request a renegotiation to raise its tariff and further restrict imports knowing that, under GATT's reciprocity rules, the domestic government would then withdraw a substantially equivalent concession that would preserve the world price and therefore deliver the tariff pair at point *A′*. Thus, while the tariff pair at point *A* is efficient, it is not robust to the type of renegotiation that GATT allows through Article XXVIII. A similar argument applies to the efficient tariff pair at point *B*, except that the roles of the foreign and domestic government are now switched. In fact, there

is only one efficient tariff pair which, if agreed to originally, would not be lost in the renegotiation process. This tariff pair is the politically optimal tariff pair, since this is the only point on the efficiency locus at which each government achieves its preferred local prices given the associated world price.

This suggests that the principle of reciprocity can be viewed as a mechanism by which governments are guided in GATT negotiations to efficient politically optimal outcomes.[12] There is a certain appeal to this finding, since the politically optimal tariffs are also those tariffs which arise when the source of the inefficiency – government's motivations to influence the terms of trade – is eliminated.

2.3. *The Importance of MFN*

If the principle of reciprocity reveals a basic economic logic, can the importance of the principle of non-discrimination in GATT practice be similarly understood? Is there a *reason* for these two principles to coexist in the same multilateral institution? This is a crucial question for the debate over regionalism versus multilateralism, since preferential agreements amount to blatant violations of the principle of non-discrimination. Hence we would like to know whether the ability of reciprocity to implement an efficient trade agreement depends in any fundamental way on the restriction that tariffs conform to MFN. We now confront this question and ask, Can reciprocity implement an efficient multilateral trade agreement that allows for discriminatory tariffs? To address this question, we must extend our two-country model to a many-country framework. This is done in Bagwell and Staiger (1997*a*), and here we simply provide an intuitive account of the results.

It is perhaps easiest to begin by imagining a multi-country world in which all tariffs conform to MFN, and consider whether reciprocity can still deliver an efficient agreement in this environment. The key observation is that, by requiring that a government levy the same tariff on imports of a good regardless of the identity of the source (i.e., exporting) country, the principle of non-discrimination preserves the simple pattern of externalities found in the two-country setting depicted above. That is, with all tariffs conforming to MFN, a common world price will prevail, and externalities continue to pass through this world price. As a consequence, reciprocity, by neutralising the world price implications of a government's tariff decisions, can guide governments to efficient politically optimal outcomes when their tariffs are non-discriminatory.

Now consider a world in which tariffs are discriminatory. If a country imports the same good from several sources and its government applies different tariff rates to imports from each source, then all else equal it would prefer to import relatively more of the good from the country to which it applies a higher tariff

[12] Actually, we establish in Bagwell and Staiger (1997*a*) that governments will indeed negotiate to politically optimal outcomes in this setting unless sufficient asymmetries are present.

owing to the lower associated world (i.e., export) price and the greater associated tariff revenue. But this implies that the pattern of externalities is now more complicated, as the government's interest in the pattern of bilateral imports from its trading partners provides a reason for concern over both world prices *and* the local prices in the markets of its trading partners (since these determine the relative shares of the import good coming from each supplying country). As a consequence, reciprocity, which serves well to neutralise externalities which travel through the world price but which is ill-equipped to handle the local-price externalities that are created when tariffs are discriminatory, fails to deliver an efficient trade agreement in this environment.

These arguments suggest a fundamental efficiency link between the principles of reciprocity and non-discrimination. Reciprocity serves to neutralise the world-price effects of a country's trade policy decisions. Provided that the externalities associated with trade intervention travel through the world price, reciprocity therefore works well as a principle with which to undo the terms-of-trade driven restrictions in trade and achieve efficient trade volumes. The principle of non-discrimination thus complements reciprocity, since it ensures that all externalities indeed are channelled through the world price.

2.4. *Why Preferential Agreements may Undermine GATT*

Having identified a basic economic logic to GATT's principle of reciprocity and an efficiency link between reciprocity and the principle of non-discrimination, we turn next to consider how preferential agreements will affect a multilateral trading system built on these two pillars. GATT's Article XXIV permits preferential agreements, provided that member countries eliminate tariffs on substantially all trade between them in a reasonable period of time. This exception to the principle of non-discrimination was controversial in its inception and has met with renewed controversy recently as many GATT members - but most especially the United States - have increasingly exercised their rights under this article to negotiate preferential trading agreements. These agreements may take either of two forms. When countries form a *free trade area*, they eliminate barriers to internal trade but maintain independent external trade policies. Under a *customs union*, member-countries also agree to harmonise their external trade policies and create a common external-tariff-setting authority.

We start with free trade areas. An immediate implication of the preceding discussion is that an efficient set of tariffs cannot be implemented under reciprocity when a free trade agreement is present. Intuitively, a free trade agreement violates MFN, and as a consequence externalities travel through both local and world prices in its presence. In this environment, as we have seen, reciprocity cannot serve to implement an efficient trade agreement. The broader implication is that the efficiency properties of a multilateral trading system founded on the principles of reciprocity and non-discrimination will be undermined when exceptions from MFN are granted for the formation of free trade areas.

Consider next the implications of customs unions for the multilateral trading system. We cannot immediately conclude from the above discussion that the introduction of a customs union is incompatible with the pursuit of an efficient set of multilateral tariffs through reciprocity. This is because the formation of a customs union creates a new environment in which, while there is now tariff discrimination, there are also fewer external-tariff-setting authorities operating, and thus it does not immediately follow from our earlier discussion that the presence of tariff discrimination will again undermine the efficiency properties of reciprocity. Nevertheless, the arguments above provide the essential intuition, once it is noted that a customs union will be analogous to a single country so long as the countries that form the union are similar in an appropriate sense. This would imply that the principle of reciprocity can then deliver an efficient agreement in the presence of a customs union between such countries so long as all *external* tariffs continue to conform to the principle of non-discrimination.

The required similarity among customs-union members that permits this result is that member-countries must share similar political goals so that the elimination of tariffs between them is internally efficient. When this relationship fails, customs unions will be like free trade areas, and it will be impossible to implement efficient tariffs under reciprocity when either type of preferential agreement is present.

More generally, we have identified a basic tension between a multilateral system built on the principle of reciprocity and the formation of preferential trading arrangements. We have shown that there is a logic to the principle of reciprocity when it is combined with the principle of non-discrimination, in the sense that the former principle serves well to deliver efficient tariffs provided that tariffs also satisfy MFN. When preferential agreements are introduced, tariffs no longer satisfy MFN, and the presumption that a multilateral system based on reciprocity will deliver an efficient outcome is severely undermined.

3. Preferential Agreements and Enforcing Multilateral Commitments

Up to this point we have compared non-cooperative outcomes with efficient trading arrangements, finding that the difference is entirely attributable to terms-of-trade externalities and that the principles of reciprocity and non-discrimination can serve to implement an efficient arrangement. We have abstracted, however, from the way in which such arrangements might be enforced. We now turn to issues of enforcement, and consider how preferential agreements can affect the ability to enforce trade commitments at the multilateral level.

3.1. *The Problem of Enforcement*

As there is no 'world jail', an international agreement must be self-enforcing if it is to be credible (see, for example, Dam, 1970), and an agreement to open markets is in turn self-enforcing only if it also specifies credible retaliatory

measures against any country that places additional restraints on trade in a way that violates the agreement. In this light, GATT can be seen as an agreement that specifies cooperative trade policies and acceptable retaliatory measures, with the maintenance of the former resting on the strength of the latter. Starting from this vantage point, the task of enforcing a reciprocal trade agreement amounts to maintaining a balance between (i) the short-term temptation to deviate unilaterally from an agreed-upon trade policy and enjoy the corresponding terms-of-trade benefits, and (ii) the long-term penalty of a consequent future loss of cooperation (i.e., the cost of a future retaliatory 'trade war').[13] Viewed in this way, it is evident that any event that alters the current temptation to cheat or the value of maintaining cooperation into the future can upset this balance, and thus that the enforceable level of cooperation may fluctuate with underlying market conditions.

A preferential agreement introduces one possible source of 'imbalance', and raises the question of whether such agreements might affect the level of multilateral cooperation that can be enforced. Various implications of preferential agreements for the ability to enforce multilateral commitments have been explored in a number of papers, including Bagwell and Staiger (1997*b*, *c* and forthcoming), Bond and Syropoulos (1996), and Bond *et al.* (1996). Here we review a number of the themes from this literature.

3.2. *Transition Effects*

Preferential agreements are typically formed over a lengthy *transition period* during which the trade policy changes associated with the agreement are being phased in, and so we begin by asking how emerging preferential agreements may affect the ability to enforce multilateral cooperation during this period of transition.[14] Our interest in this question is due in part to historical and current experiences with regard to preferential agreements and multilateral tariff cooperation. Beginning in 1957, the EC customs union was formed over a twelve year phase-in period, and it underwent a period of major expansion to include the United Kingdom, Denmark and Ireland beginning in 1972. These episodes of customs union formation and enlargement corresponded with periods of enhanced multilateral tariff cooperation and, as a recent WTO report concludes (WTO, 1995b, pp. 53–4), were factors behind the launching of the GATT Dillon Round (1960–2), Kennedy Round (1964–7) and Tokyo Round (1973–9) of multilateral negotiations. More recently, important prefer-

[13] We draw a distinction between *unilateral deviations* from an agreed-upon trade policy and the lawful *withdrawal or modification of a previously negotiated concession* under Article XXVIII. The former may go undetected for some time but, once observed by trading partners, would trigger a retaliatory 'trade war'. The latter must be pre-announced to trading partners who are then free to simultaneously withdraw substantially equivalent concessions under the procedures of Article XXVIII described in Section 2.2 above.

[14] GATT's Article XXIV acknowledges the practical need for 'interim agreements' to facilitate the process of preferential integration, and only requires that the transition to a completed free trade area or customs union be accomplished 'within a reasonable length of time'.

ential agreements include the 1988 U.S.-Canada free trade agreement and its expansion to include Mexico in the NAFTA. The implementation of these agreements, by contrast, appears to have taken place against a backdrop of strained multilateral relations in which preferential initiatives are viewed as a potential threat to the GATT system.[15]

In Bagwell and Staiger (1997*b*, *c*), we present formal models broadly consistent with these observations. To interpret the findings of these papers, we note that there are two principal effects of preferential agreements that are crucial in determining how they will affect enforcement at the multilateral level: a *trade diversion effect*, under which trade volumes among member countries increase at the expense of trade between member and non-member countries; and a *market power effect*, which occurs if the member countries form a customs union and adopt a common external tariff policy that enables them to impose higher tariffs on their multilateral trading partners should such punitive tariff action be desired.

Consider first the transition to a free trade area. The main ideas can be understood in terms of a three-country setting. Once countries *A* and *B* become firmly engaged in the lengthy transition process that will culminate in a free trade area, country *C* faces the prospect that it is currently trading more extensively with country *A* than it is likely to in the future, since more of country *A*'s future trade will be diverted to its free trade partner once the free trade agreement is fully implemented. Thus, while country *C*'s current temptation to exploit its power over the terms of trade with country *A* is largely unaffected, owing to its as yet undiminished current trade volume with country *A*, country *C* is no longer as fearful of a future trade war with country *A*, since it expects that it will in any case trade less with country *A* in the future. Incentives are thus temporarily thrown out of balance as the transition to a free trade area between countries *A* and *B* begins, and as a result the trade policies that countries *A* and *C* can enforce will be less cooperative during the associated transition phase. It follows that the transition to a free trade area will bring about a period of temporarily heightened multilateral trade tensions in which trade disputes proliferate and further efforts to reduce multilateral tariffs become temporarily stalled. These predictions seem broadly compatible with recent experiences.

On the other hand, when countries *A* and *B* are in the lengthy transition process that will culminate in a customs union, country *C* is faced with the emergence of a new 'market power' effect. Intuitively, when a customs

[15] According to the WTO report, existing preferential agreements were a less significant factor in the 1986 launching of the Uruguay Round of GATT negotiations, because '... at the time, regional integration was still confined mainly to Western Europe, with the United States maintaining its traditional multilateralism.' (WTO, 1995b, p. 54). Nevertheless, while the failure of these negotiations to conclude at the Brussels Ministerial in December 1990 reflected the strained mulitlateral relations of the time, this failure together with the subsequent increase in new preferential initiatives after 1990 were '... major factors in eliciting the concessions needed to conclude the Uruguay Round' in 1994 (WTO, 1995b, p. 54), as they raised the specter that a failed Uruguay Round would lead to a world in which future trade and economic relations would be based primarily upon preferential agreements.

union is formed, its common external tariffs enable it to exert greater power over the terms of trade. As a consequence the union will find high import tariffs more tempting than do its individual member countries prior to uniting. This suggests that a trade war initiated in the transition phase might have heightened negative consequences for country *C* once the union is formed, as countries *A* and *B* will retaliate even more aggressively once they select a common tariff. Accordingly, country *C*'s current temptation to cheat in the transition phase is now more than outweighed by its fear of the retaliation that a customs union could later mete out. Country *C*'s incentives are again temporarily out of balance, but in this case it will tolerate even more liberal multilateral tariffs before the difficulty of enforcement again begins to bind. In this way, the transition to a customs union involves a period of improved multilateral tariff cooperation, much as historical experience suggests.

3.3. *Steady State Effects*

Finally, in a more recent paper (Bagwell and Staiger, forthcoming), we ignore the transitional effects of preferential agreements and consider instead a three-country world in which two countries are better at cooperating than is the third. In particular, we assume that countries *A* and *B* are more patient, and are thus better able to enforce liberal trading policies, than is country *C*. Within this setting, we compare two trading regimes.

Consider first the case in which countries are constrained to adopt MFN tariffs. Countries *A* and *B* then cooperate best by acting as 'hegemons', extending tariff cuts to country *C* that exceed the cuts that country *C* offers in return. The impatient country is therefore 'pooled in' with the patient countries under MFN, and it gets to free ride on their liberalisation efforts.

Consider next a second regime, in which countries *A* and *B* form a preferential agreement, enabling them to offer a tariff to country *C* that differs from the zero tariff that they extend to one another. With the ability to discriminate, the patient countries need no longer cooperate multilaterally to cooperate bilaterally, and so the impatient country loses its free rider benefits.

The discriminatory tariff that is offered to country *C* in this second trading regime will often exceed the tariff it would face were the preferential agreement not allowed, and this leads to an overall deterioration in multilateral tariff cooperation once the preferential agreement is formed. However, the opposite can also occur, as it is possible that the discriminatory tariff offered to country *C* falls with the introduction of the preferential agreement between countries *A* and *B*. Which outcome occurs, and hence whether preferential agreements act as 'stumbling blocs' or 'building blocs' for multilateral tariff cooperation, depends in this setting on how close the multilateral system can get to an efficient trade agreement in the absence of tariff discrimination, which depends in turn on how patient the two 'hegemons' are. Weighing these factors, we find that preferential agreements can facilitate multilateral liberalisation, and that they have their most desirable effects on the multilateral

system precisely when multilateral enforcement mechanisms are ineffective and the multilateral system is working poorly.

4. Is Bilateralism Bad?

We have argued that the multilateral trading system can be understood as a cooperative arrangement among governments that is designed to eliminate the inefficient trade restrictions that are associated with governments' ability to manipulate the terms of trade. In the absence of enforcement difficulties, we find that GATT's principles of reciprocity and non-discrimination can work in tandem to implement an efficient outcome. As a corollary, we find that reciprocity cannot serve to implement an efficient agreement in the presence of free trade areas. An efficient agreement can be implemented under reciprocity in the presence of a customs union, but only if the union members have similar political preferences. These conditions are quite stringent, and so we offer little support for the hypothesis that reciprocity can deliver an efficient multilateral trade agreement in the presence of preferential trade agreements. Instead our results offer support for the view that preferential agreements pose a threat to the existing multilateral system.

Enforcement concerns, however, should not be ignored, as the threat of future retaliation may not be sufficient to deliver a fully efficient multilateral agreement. Consequently, significant changes in the trading environment, such as occur when major preferential integration initiatives are undertaken, can have an impact on the level of multilateral cooperation that can be enforced. This impact will depend critically on the period of analysis, i.e., transition or steady state, on the form that the preferential agreement takes, i.e., free trade agreement or customs union, and on the strength of the multilateral enforcement mechanism. Nevertheless, these results serve to qualify the more negative view of preferential agreements that obtains in the absence of serious limitations to the multilateral enforcement mechanisms, and the implied qualifications achieve their greatest force when multilateral enforcement mechanisms are at their weakest.

Finally, as to the question we posed at the outset, our results suggest that the efficiency of the multilateral trading system will be compromised by the creation of preferential agreements unless multilateral enforcement mechanisms are sufficiently weak. In this light, further strengthening of the enforcement mechanisms of the GATT/WTO will undercut the case for preferential agreements.

Columbia University

University of Wisconsin-Madison

References

Bagwell, Kyle, and Staiger, Robert W. (1996). 'Reciprocal trade liberalization,' NBER Working Paper no. 5488, March.

Bagwell, Kyle, and Staiger, Robert W. (1997*a*). 'An economic theory of GATT,' NBER Working Paper no. 6049, May.

Bagwell, Kyle, and Staiger, Robert W. (1997*b*). 'Multilateral tariff cooperation during the formation of customs unions,' *Journal of International Economics*, February.
Bagwell, Kyle, and Staiger, Robert W. (1997*c*). 'Multilateral tariff cooperation during the formation of free trade areas,' *International Economic Review*, May.
Bagwell, Kyle, and Staiger, Robert W. (1998). 'Regionalism and multilateral tariff cooperation,' in (John Piggott and Alan Woodland, eds.), *International Trade Policy and the Pacific Rim*, London: Macmillan, forthcoming.
Baldwin, Richard, (1987) 'Politically realistic objective functions and trade policy,' *Economic Letters*, vol. 24.
Bhagwati, Jagdish, (1991). *The World Trading System at Risk.* Princeton, New Jersey: Princeton University Press.
Bond, Eric W., and Syropoulos, Costas (1996). 'Trading blocs and the sustainability of inter-regional cooperation,' in M., Canzoneri, W., Ethier, and V. Grilli, (eds.) *The New Transatlantic Economy*, Cambridge: Cambridge University Press.
Bond, Eric W., Syropoulos, Costas and Winters, Alan A., (1996). 'Deepening of regional integration and external trade relations,' CEPR Discussion Paper No. 1317.
Dam, Kenneth W., (1970). *The GATT: Law and International Economic Organization*, Chicago: University of Chicago Press.
Enders, Alice, (1997). 'The origin, nature and limitations of reciprocity in GATT 1947,' mimeo. Geneva: WTO, October.
Feenstra, Robert C., (1989). 'Symmetric pass-through of tariffs and exchange rates under imperfect competition: an empirical test,' *Journal of International Economics*, vol. 27, pp. 25–45.
Feenstra, Robert C., (1995). 'Estimating the effects of trade policy,' in (Gene M. Grossman and Kenneth Rogoff, eds.) *The Handbook of International Economics*, vol. 3, North Holland.
Gros, D., (1987) 'A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade,' *Journal of International Economics*, vol. 23, pp. 357–67.
Jackson, John H., (1989). *The World Trading System.* Cambridge: The MIT Press.
Johnson, Harry G., (1953–4). 'Optimum tariffs and retaliation,' *Review of Economic Studies*, vol. 21, no. 2.
Kennan, John and Riezman, Raymond (1988). Do big countries win tariff wars?,' *International Economic Review*, vol. 29, pp. 81–5.
Krugman, Paul R., (1991*a*). 'Is bilateralism bad?' in (Elhanan Helpman and Assaf Razin, eds.) *International Trade and Trade Policy*, Cambridge, MA and London, England. The MIT Press, pp. 9–23.
Krugman, Paul R., (1991*b*). '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.
Krugman, Paul R., (1997). 'What should trade negotiators negotiate about?,' *Journal of Economic Literature*, vol. 35, March 1997, pp. 113–20.
Maggi, Giovanni, and Rodriguez-Clare, Andres (1998). 'The value of trade agreements in the presence of political pressures,' *Journal of Political Economy*, forthcoming.
Mayer, Wolfgang, (1981). 'Theoretical considerations on negotiated tariff adjustments,' *Oxford Economic Papers*, vol. 33, pp. 135–53.
Mill, John Stewart, (1844). *Essays on Some Unsettled Questions of Political Economy*, London: Parker.
Scitovszky, Tibor, (1942). 'A reconsideration of the theory of tariffs,' *Review of Economic Studies*, vol. 9.
Staiger, Robert W., (1995). 'International rules and institutions for trade policy,' in (Gene M. Grossman and Kenneth Rogof, eds.) *The Handbook of International Economics*, vol. 3, North Holland.
Torrens, Robert, (1844), *The Budget: On Commercial and Colonial Policy.* London: Smith, Elder.
WTO, (1995a). *Analytical Index: Guide to GATT Law and Practice*, vol. 2. Geneva: WTO, April.
WTO, (1995b). *Regionalism and the World Trading System.* Geneva: WTO, April.