The Role of Multilateral Institutions in International Trade Cooperation

By GIOVANNI MAGGI*

The World Trade Organization (WTO) lacks the power to directly enforce agreements. It is therefore important to understand what role the WTO can play to facilitate international cooperation, and whether a multilateral institution can offer distinct advantages over a web of bilateral agreements. This paper examines two potential benefits of a multilateral trade institution: first, verifying violations of the agreements and informing third parties, thus facilitating multilateral reputation mechanisms; second, promoting multilateral trade negotiations rather than a web of bilateral negotiations. The model suggests that a multilateral approach is particularly important when there are strong imbalances in bilateral trading relationships. (*JEL* F13)

Since 1947, trade barriers between nations have decreased dramatically; for example, the average ad valorem tariff on industrial goods has declined from about 40 percent to less than 4 percent. Does the success of trade liberalization efforts have something to do with the presence of the WTO (formerly known as the General Agreement on Tariffs and Trade, GATT)? If so, what role has the WTO played in facilitating trade cooperation, given that it has no direct enforcement power?

The trade literature has only recently started to address this question. Avinash Dixit (1987), Kyle Bagwell and Robert W. Staiger (1990), and Raymond Riezman (1991), among others, examine issues of trade cooperation by way of two-country models in which governments choose trade policies in a repeated-game setting. In these models there is typically a great multiplicity of equilibria, including very inefficient ones. These models suggest that the WTO may help countries coordinate on more efficient equilibria.

A second potential role of the WTO, and specifically of its Dispute Settlement Procedure (DSP), is suggested by Thomas Hungerford (1991) and Dan Kovenoch and Marie Thursby (1993): the DSP may act as an information-gathering agency that is able to discern between true violations of the agreement and mistaken perceptions, thus facilitating the use of a bilateral reputation mechanism to support cooperation.[1] The emphasis here is on bilateral monitoring: the idea is that the DSP can improve monitoring by the country directly affected by the trade policy.[2]

In the present paper I focus on two other roles that the WTO-DSP can perform, both of

* Department of Economics, Princeton University, Princeton, NJ 08544 (e-mail: maggi@princeton.edu). I acknowledge financial support from the Alfred P. Sloan Foundation dissertation fellowship. I would like to thank Kyle Bagwell, Serdar Dinc, Avinash Dixit, Faruk Gul, Avner Greif, Thomas Hellmann, Paul Milgrom, Andres Rodriguez-Clare, Gabriel Srour, Barry Weingast, Frank Wolak, Beth Yarbrough, and the participants in various seminars for various comments and discussions. I thank Nina Pavcnik and Caroline Thompson for research assistance. I am particularly indebted to Robert Staiger, for his invaluable advice, and to Gene Grossman, who provided detailed comments on two versions of this paper. I also thank two anonymous referees for very helpful suggestions. The paper was revised while I was visiting the Economics Department at NYU, which I thank for the hospitality.

[1] Hungerford (1991) suggests this potential role of the DSP but does not formalize it. In fact, in his model the DSP can only play a negative role, because the DSP investigations are assumed to be uninformative and costly.

[2] Kovenoch and Thursby (1993) suggest also another possibility: the existence of the DSP may instill in countries a sense of "international obligation" that increases the cost of violating the agreements beyond any retaliation that a violation may trigger from trading partners.

which are intrinsic to its multilateral nature. First, it can verify violations of the agreement and inform third countries, thus facilitating multilateral enforcement efforts. Second, it can promote a multilateral rule-making procedure in place of a web of bilateral negotiations. Since the papers mentioned above all utilize two-country models, they cannot examine these aspects.

To start with the informational role, the idea is the following. Suppose country A commits a violation against country B, and this violation is observed by country B but not by the rest of the trading community. A potential role for the DSP in this case is to identify the violation and bring it to the attention of third countries, exposing the offending country to a loss of reputation in the trading community. Notice the difference with respect to the bilateral-monitoring role discussed above: the idea here is that the institution can improve monitoring by third countries, rather than by the "second" country.[3]

Two questions arise about the relevance of this argument. First, how important is the scope for this kind of information dissemination for trade policies? Some trade policies, such as tariffs and quotas, are fairly transparent, however there are two important dimensions in which trade policies are much less transparent, particularly to third countries: (i) There are a number of nontariff barriers to trade (such as government practices that favor domestic producers over foreign ones, antidumping actions, technical and safety requirements on imported products, etc.) that can be hard for third countries to perceive. (ii) Current trade agreements allow countries to increase tariffs or quotas above their "baseline" levels under certain contingencies.[4] If a country erects a trade barrier by invoking such special circumstances, third countries may not know what the underlying circumstances are, hence whether there has been a violation of the agreement.

The second issue concerns the idea that an offending country can be punished by a loss of cooperation with third countries. What forms can this loss take in reality? There are several ways in which the WTO community can respond to a violation of the agreements. First, WTO members can withdraw trade concessions to the defecting country. For example, it is the interpretation of several scholars that Article XXIII of the GATT agreement provides for the possibility of expelling a repeatedly offending country from GATT (see John H. Jackson, 1969 pp. 186–87; Michael Finger, 1993).[5] Second, WTO members can impose costs on the offending government in more subtle ways, by withdrawing some of their "goodwill" toward that government: they can be less forthcoming with the offending country in subsequent negotiations, in the same or related areas of cooperation; they can be more reluctant to enter new agreements with the offending country; or, if trade liberalization takes place in a gradual fashion, they can slow down the liberalization process vis-à-vis the offending country. Also, third countries can reduce their cooperation with the offending country at the level of institutional procedures; for example, if country A does not follow the recommendations of the DSP panel, third countries may feel free to do the same in future disputes against country A.[6]

[3] The idea that institutions without enforcement power can serve to complement reputation mechanisms by disseminating information has already been formalized by Paul Milgrom et al. (1990) and Avner Greif et al. (1994) in the context of medieval trading institutions. These works differ from the present paper both in their formal structure and in the issues they examine. The role of multilateral reputation mechanisms is also explored in Michihiro Kandori (1992), who analyzes a game in which traders match randomly period by period, and in Jonathan Bendor and Dilip Mookherjee (1990), who analyze a multilateral version of the repeated prisoner's dilemma game; I will refer again to this paper in footnote 16.

[4] One example is Article XIX of GATT, which allows special trade protection in case of an unexpected import surge that threatens to injure a domestic industry.

[5] One historical case in which severe multilateral sanctions were explicitly threatened against a single country is the dispute over the U.S. dairy quotas in 1951: in response to the illegal imposition of U.S. import quotas on dairy products coming mostly from the Netherlands, "the Contracting Parties had brought out their biggest guns against the dominant partner. They had threatened everything that could be threatened, including the collapse of the Agreement itself." (Robert Hudec, 1975 p. 167.)

[6] The view of the WTO-DSP as performing an information-dissemination role is in line with the writings of many scholars in international law and international

Having discussed the kinds of punishment that the WTO community *potentially* can inflict on the offender, there are two questions that can be addressed: one normative (what the WTO community should do to the offender), the other positive (what the WTO community does do to the offender, and why). The theory presented in this paper, which I will preview shortly, can be interpreted as addressing the normative question. Whether multilateral punishment threats have played a role in practice—and thus, whether the theory can be interpreted as explaining the actual role of the WTO—is a more subtle question. First, to paraphrase Thomas Schelling's expression, the effectiveness of an army sometimes must be judged by how little it is used. While it is true that explicit multilateral sanctions have never been observed in the WTO, it is also true that there have been no cases of blatant and repeated violations of key WTO rules, even by strong countries against vulnerable trading partners; it is reasonable to think that strong countries have been deterred from abusing weaker partners by an implicit threat of multilateral sanctions.[7] Second, more subtle forms of multilateral punishments—such as the ones discussed in the previous paragraph—are hard to pinpoint in practice, even if they do take place. The best we can hope for is some indirect, anecdotal evidence from the history of GATT disputes. In the last section of the paper, I review a number of disputes which share the following theme: a strong country was accused by a weaker country of violating an agreement, and the strong disputant ended up withdrawing the contested policy, in spite of a clear opportunity to abuse the weaker partner. These anecdotes, I will argue, are suggestive of an effective—if subtle—role of multilateral enforcement pressures in improving compliance with international trade agreements.

With regard to the role of multilateral punishment threats, it is useful to anticipate a result of the model that is important for evaluating the positive relevance of the theory. While the model shows that a multilateral enforcement mechanism is desirable, it also points out that third-party sanctions should be "minimal," in a sense to be made precise, and should be threatened only for certain violations, namely those that are hard to deter with bilateral sanctions alone. This result is broadly consistent with the fact that, in the GATT experience, multilateral enforcement pressures seem to play a more subtle and selective role than bilateral ones.

In what follows, I give a brief overview of the model and of the main results. The theoretical analysis is based on a multicountry model in which governments repeatedly select import barriers. To examine the potential benefits of an institution that verifies and publicizes violations, I compare two extreme scenarios: a WTO-less world, in which third countries remain unaware of violations, and a world with WTO, in which all governments are informed of any violations. In this setting, the only possible benefit of information dissemination is that it enables third countries to punish violators. Thus, the question becomes

relations. Jock Finlayson and Mark Zacher (1981 p. 587) write: "Observers of the GATT often argue that improving the quality and quantity of information about international trade policy has been one of the regime's major contributions. More importantly, without the provision of data and information concerning members' trade policies, behavior could not be effectively monitored and therefore the ability to implement regime rules would suffer."

Judith Goldstein (1993) writes: "The GATT is among the most successful of the multilateral organizations built at the close of World War II. ... The virtues of reciprocity norms, of precise monitoring procedures, and of the dissemination of information have been cited as central elements in efficient international organizations." Other international-relations scholars who have emphasized the informational role of international institutions are Stephen Krasner (1983), Robert Keohane (1984), Kenneth Oye (1986), Beth Yarbrough and Robert Yarbrough (1992), and Lisa Martin (1993).

[7] GATT commentators often argue that countries are deterred from violating trade agreements not just by the prospect of bilateral retaliation, but by the fear that the whole trading system may unravel as a consequence, or in other words, by the fear of a multilateral breakdown of cooperation. For example, referring to a speech by the ex-Director-General of GATT, Arthur Dunkel, John Croome (1995 pp. 11–12) writes: "Dunkel ... concluded that governments are being restrained from a substantial slippage towards protectionism only by 'a kind of balance of terror': a fear that if they resorted to trade restrictions these would evoke retaliation, as well as undermining the trading system as a whole." I thank Kyle Bagwell for providing me with this reference.

whether a multilateral enforcement mechanism can sustain more cooperative outcomes than a bilateral one.[8]

A key feature of the model is the presence of bilateral imbalances of power, where the "more powerful" country in a given pair is the one that stands to lose less (or to gain) from a trade war. In this situation, a multilateral enforcement mechanism may be beneficial, because it allows a transfer of enforcement power across relationships that is not possible under bilateral enforcement. In particular, the optimal self-enforcing agreement requires stronger countries to make more generous concessions than weaker countries, but this can be implemented only with a multilateral enforcement mechanism. The gains from multilateral enforcement do not arise simply because punishments are more severe, but specifically because of imbalances in power: in the benchmark case where power is balanced, there are no benefits from multilateralism. The analysis also indicates that international monetary transfers, if feasible, can mitigate the effects of bilateral power imbalances, but cannot fully substitute for multilateral enforcement. While the model examined in this paper focuses on imbalances in power, these are not the only factor that gives rise to gains from multilateral enforcement; in Section IV, I will discuss other factors that can make multilateral sanctions desirable.

The next part of the analysis takes a closer look at the optimal enforcement mechanism. The key finding is that third-party sanctions should be kept within limits. In particular, the analysis reveals that: (i) Past a point, increasing the severity of third-party sanctions in response to a bilateral violation of the agreement will not enhance cooperation; in this sense, "maximal" third-party sanctions are not necessary. Also, the threat of third-party sanctions is necessary only for certain violations, namely those by stronger countries against weaker countries. (ii) If the monitoring of trade policies is slightly imperfect, and if inflicting a punishment is slightly more costly for third parties than for second parties, then third-party sanctions should be *minimal,* in the sense of being the least severe that achieve all gains from multilateral enforcement.

The model is then extended to examine a situation in which the DSP again is called upon to verify and publicize violations, but where now a government can block this process by, say, refusing to cooperate with the investigation. In this situation, governments must be induced to comply not only with substantive rules, but also with procedural rules (e.g., facilitating the DSP operations). The interesting question is whether the governments' power to block the transmission of information undermines the multilateral enforcement mechanism. The analysis suggests that the nonenforceability of procedural rules does not impose a binding constraint on the enforcement system, as long as the DSP can publicize procedural violations.

Aside from enforcement issues, the model indicates that a multilateral approach may be important also at the level of trade negotiations. Trade negotiations are modelled as a process of Nash bargaining over the set of self-enforcing trade policy configurations. It is found that, in the presence of power imbalances, multilateral bargaining enables countries to achieve deeper trade liberalization than a web of bilateral negotiations. The interesting aspect of this result is that each bilateral negotiation is locally efficient (i.e., it is efficient conditional on the outcome of other bilateral negotiations), and there are no externalities across bilateral relationships, yet a system of bilateral negotiations is globally inefficient. The inefficiency of bilateral bargaining is similar to the market failure caused by incomplete markets; if one regards trade negotiations as a market where countries exchange trade concessions, bilateral bargaining is inefficient because the market is segmented. At a broad level, this result suggests a further potential role of trade institutions such as the WTO: to the extent that they can promote multilateral rule-making procedures, they can prevent distortions arising from bilateral imbalances of power.

The paper is structured as follows. Section I presents the basic model. Section II examines the implications of multilateralism at the enforcement level. Section III discusses the

[8] I will use the expressions "reputation mechanism" and "enforcement mechanism" interchangeably.

efficiency gains from multilateral trade negotiations. Section IV discusses some extensions of the model. Section V offers concluding remarks.

I. The Basic Model

I consider a three-country trading system in which bilateral relationships are separable, in the sense that trade policies in a given bilateral relationship do not affect trade flows in the other two bilateral relationships. In this setting, trade policies do not cause trade diversion. I will argue later that the possibility of trade diversion tends to strengthen the paper's main results on the gains from multilateral enforcement and rule-making.

A simple way to rule out trade diversion is to assume that each pair of countries trade two distinct goods that are neither supplied nor demanded by the third country. More formally, let $\mathcal{N} = \{A, B, C\}$ denote the set of countries and $\mathcal{N}\backslash i$ the pair of countries that does not include country i. The numeraire good is indicated by a "0" subscript. Each nonnumeraire good is indexed by an ordered pair (i, j), with $i \in \mathcal{N}, j \in \mathcal{N}, i \neq j$. Country i's representative consumer has the following utility: $U^i = x_0 + \Sigma_{j \in \mathcal{N}\backslash i} [u(x_{ij}) + u(x_{ji})]$, where x_{ij} denotes consumption of good (i, j). Country i is endowed with z_0 units of the numeraire good and with z_{ij} units of good (i, j), $j \in \mathcal{N}\backslash i$.[9] Given this structure of preferences and endowments, good (i, j) is exported from country i to country j under free trade. To be more concrete, consider for example country A: this country is endowed with goods (A, B) and (A, C), which it exports respectively to countries B and C, and imports goods (B, A) and (C, A) respectively from country B and C; in country A there is no demand for good (B, C) or good (C, B).

Let p^x_{ij} and p^m_{ij} denote the price of good (i, j) respectively in country i (the exporting country), and in country j (the importing country). Governments choose specific import tariffs on the nonnumeraire goods (the numeraire good is assumed to be freely traded). I focus on tariffs because they are particularly simple to model; however, the qualitative results would be the same if governments chose nontariff import barriers, as long as the game has the structure of a prisoner's dilemma. An import tariff creates a wedge between the local price (p^m_{ij}) and the offshore price (p^x_{ij}). Thus, if τ_{ij} denotes the import tariff on good (i, j), the following relationship holds, provided the tariff is not prohibitive:

$$p^m_{ij} = p^x_{ij} + \tau_{ij}. \tag{1}$$

Export taxes and import subsidies are not allowed; I will discuss the role of this assumption later in this section.

The population in each country is a continuum of individuals whose total size is normalized to one. Country i's local consumers demand good (i, j) in amount $d(p^x_{ij})$ [where $d(\cdot)$ is the inverse of $u'(\cdot)$] and good (j, i) in amount $d(p^m_{ji})$. Consumer surplus in country i is given by $\Sigma_{j \in \mathcal{N}\backslash i} [s(p^m_{ji}) + s(p^x_{ij})]$, where $s(p) = u[d(p)] - pd(p)$. The function $u(\cdot)$ is assumed to be quadratic, so that $d(\cdot)$ is linear.

Market clearing for good (i, j) requires:

$$d(p^x_{ij}) + d(p^m_{ij}) = z_{ij}. \tag{2}$$

Equations (1) and (2) implicitly define the market-clearing prices $p^x_{ij}(\tau_{ij}; z_{ij})$ and $p^m_{ij}(\tau_{ij}; z_{ij})$. It is direct to verify that:

$$\partial p^x_{ij}/\partial \tau_{ij} < 0, \partial p^m_{ij}/\partial \tau_{ij} > 0, \tag{3}$$

$$\partial p^x_{ij}/\partial z_{ij} < 0, \partial p^m_{ij}/\partial z_{ij} < 0.$$

The volume of country i's imports from country j is given by $m(\tau_{ji}; z_{ji}) \equiv d(p^m_{ji}(\tau_{ji}; z_{ji}))$, and the associated tariff revenue is given by $r(\tau_{ji}; z_{ji}) \equiv \tau_{ji} m_{ji}(\tau_{ji}; z_{ji})$. The owners of good (i, j) earn profits in amount $\pi(\tau_{ij}; z_{ij}) \equiv p^x_{ij}(\tau_{ij}; z_{ij}) z_{ij}$. It is convenient to define the per-good consumer surplus for imported and exported goods, respectively as $s^m(\tau_{ji}; z_{ji}) \equiv s(p^m_{ji}(\tau_{ji}; z_{ji}))$ and $s^x(\tau_{ij}; z_{ij}) \equiv s(p^x_{ij}(\tau_{ij}; z_{ij}))$.

[9] The endowment of the numeraire good is assumed to be large enough that it is always consumed in positive amounts by each agent. Under this condition, the marginal utility of income is fixed, thus the market for each nonnumeraire good can be analyzed in partial-equilibrium fashion. Trade in the numeraire good is then determined residually by the condition of overall trade balance for each country.

Also let $U_m(\tau_{ji}; z_{ji}) \equiv s^m(\tau_{ji}; z_{ji}) + r(\tau_{ji}; z_{ji})$ and $U_x(\tau_{ij}; z_{ij}) \equiv s^x(\tau_{ij}; z_{ij}) + \pi(\tau_{ij}; z_{ij})$. This notation will bring out more clearly the separable structure of payoffs.

Government i is assumed to maximize its own citizens' welfare. To derive an expression for welfare, notice that the indirect utility of an individual with income y is given by $y + \Sigma_{j \in \mathcal{N} \setminus i} [s(p_{ji}^m) + s(p_{ij}^x)]$. Aggregate national welfare is given by the sum of indirect utilities over the population. Since population size equals one, and aggregate income is the sum of profits and tariff revenue, welfare is given by the sum of profits, tariff revenue, and consumer surplus. The welfare function can be expressed in the following way:

$$(4)\quad W^i = z_0 + \sum_{j \in \mathcal{N} \setminus i} [U_m(\tau_{ji}; z_{ji}) + U_x(\tau_{ij}; z_{ij})].$$

Having described the governments' payoffs, I now turn to the analysis of the static trade policy game. Governments choose import barriers simultaneously. Due to the separability of payoffs, country i's best-response tariff on good j is independent of all other tariffs, and given by $\text{argmax}_\tau U_m(\tau; z_{ji}) \equiv \tau^N(z_{ji})$. It is direct to verify that the optimal tariff is strictly positive, as a small tariff generates a second-order loss in consumer surplus and a first-order improvement in the country's terms of trade: this is the usual beggar-thy-neighbor motive for a tariff, which was formalized by Harry G. Johnson (1953–1954). Clearly, the one-shot game has a unique Nash equilibrium, in which country i's tariff on good j is given by $\tau_{ij} = \tau^N(z_{ji})$.

At this point I impose a symmetry restriction on the game. Countries are symmetric, but each has one bilateral relationship in which it is a net exporter, and one in which it is a net importer. Think of countries as located at the vertices of a triangle, and assume that the representative country is characterized by endowment levels (z^L, z^R), where z^L (respectively z^R) is the endowment of the good exported to the country's left (respectively right) trading partner. To fix ideas, suppose $z^L \geq z^R$, so that the representative country is a net importer from the right partner and a net exporter to the left partner. At the Nash equilibrium, each country selects the same tariffs, $\tau^{NL} \equiv \tau^N(z^R)$ and $\tau^{NR} \equiv \tau^N(z^L)$, on imports coming respectively from the left and the right partner. The tariff imposed on the right partner is higher ($\tau^{NR} > \tau^{NL}$), as a higher volume of imports gives the government a higher incentive to tax them. Figure 1 illustrates the situation.

The trade-policy game has the structure of a multilateral prisoner's dilemma, as all countries would be better off under global free trade than at the Nash equilibrium. At the same time, the system is characterized by *bilateral imbalances of power,* in the sense that the partners in each bilateral relationship stand to lose different amounts from a trade war. The more "powerful" country in a given bilateral relationship is defined as the one that loses less (or gains) if the two countries move from free trade to the static Nash tariffs. In this setting, each country is more powerful than its right partner, and weaker than its left partner.[10] An important benchmark is the one in which the trading system is symmetric not only globally but also in every bilateral relationship (i.e., $z^L = z^R$). Comparison with this benchmark will be useful to understand the specific implications of imbalances in power.

In this model, bilateral imbalances in power are captured in a very stylized way: the presence of trade imbalances, together with the assumption of no export taxes, implies that the net importer is more powerful than the net exporter. One can think of many other reasons why different countries may stand to lose different amounts from a trade war; the assumption of no export taxes should not be regarded as restrictive, but rather as a simple way to illustrate a more general point about the implications of unequal power.[11]

[10] In Section IV I will discuss a variant of the model in which the trading system is characterized by global power imbalances.

[11] The assumption of no import subsidies also deserves some discussion. If import subsidies were allowed, they would not be used at the Nash equilibrium of the one-shot game, thus they would not affect power imbalances as they are defined in the text. However, they could be used in the repeated game to make bilateral enforcement more effective, because they imply a transfer from importer to exporter (as will be made clear in Section II, subsection A, transfers can improve the effectiveness of bilateral enforcement, although they cannot fully substitute for

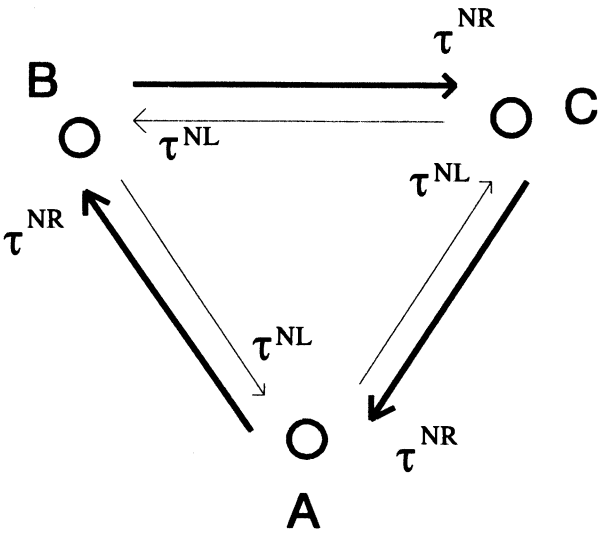

FIGURE 1. THE STATIC NASH EQUILIBRIUM

Another simplifying assumption of the model is that governments maximize national welfare. The model could be extended to include political motives in the governments' objectives, as for example in Richard E. Baldwin (1987). As long as trade barriers impose an externality on trading partners, the game will have the structure of a prisoner's dilemma game (as pointed out by Bagwell and Staiger, 1999). The only difference would be that the fully cooperative outcome may no longer be free trade, but this would not affect the results about the benefits of a multilateral approach to cooperation relative to a bilateral one.

II. Gains from Multilateral Enforcement

I now turn to the analysis of the supergame, i.e., the infinite repetition of the game just described. In the supergame, governments have the possibility to sustain freer trade than in the one-shot game, by threatening to revert to high protection levels in case of defections. Let $\delta \in (0, 1)$ denote the discount factor between periods.[12] The main objective of this section is to examine the potential role of an institution that verifies and publicizes defections. For this purpose, the institution is regarded as a mechanism that disseminates information automatically, honestly, and without delay. I will compare two stylized scenarios: (1) a world without such an institution, in which only bilateral histories are observed (I will sometimes refer to this as a world with "bilateral monitoring"); and (2) a world with such an institution, in which the entire history is common knowledge (one with "multilateral monitor-

multilateral enforcement). At the same time, it should be noted that import subsidies are dominated by lump-sum transfers, because they are an inefficient way to transfer income; if transfers are feasible, import subsidies are redundant. Thus, if import subsidies were introduced in the extended model of Section II, subsection A, its results would not be affected.

[12] The infinitely repeated game with discount factor δ can be interpreted as a repeated game in which each government faces a hazard rate (constant over time and across governments) that the game will continue. In this case, $\delta = he^{-rL}$, where h is the hazard rate, r incorporates the rate of time preference, and L is the length of the period. Under perfect monitoring, L can be interpreted as the length of time it takes to observe the trading partners' policies and respond to them. The reader is referred to the discussion in Staiger (1995 pp. 1520–21), who argues convincingly that the response lag for trade policies is often lengthy, so we should not expect δ to be very close to one in reality.

ing''). These informational structures are extreme, but they allow one to focus sharply on the potential advantages of information dissemination in the trading system. The question is whether countries can sustain more cooperative outcomes under multilateral monitoring than under bilateral monitoring. Since the only possible benefit of multilateral monitoring in this setting is to make third-party sanctions possible, the exercise boils down to asking whether third-party sanctions can facilitate cooperation.

I will focus on sequential equilibria in which each country selects the same import barriers (τ^L, τ^R) at all times along the equilibrium path. As for the players' behavior following a deviation, I consider two kinds of punishment strategies. Define a bilateral enforcement mechanism (BEM) as a punishment strategy whereby a defection by one country against another is followed by a permanent reversion of (only) these two countries to the static Nash tariffs, and a multilateral enforcement mechanism (MEM) as a punishment strategy whereby any defection is followed by permanent Nash reversion in both bilateral relationships which the defector is involved in. Clearly, both of these punishment strategies are credible, as no government has an incentive to deviate from the static Nash tariffs.[13] The punishment inflicted by third countries here takes an extreme form, but the reader should keep in mind that less severe third-party sanctions are sufficient to realize all gains from multilateral enforcement, as will be shown in Section II, subsection B.

I will compare the maximum symmetric equilibrium payoffs sustainable under BEM and MEM. For a given enforcement mechanism, I will refer to the tariff pair that maximizes the symmetric equilibrium payoffs as the ''most cooperative'' tariff pair.

The first step is to determine the most cooperative tariffs under BEM. Let $U^L(\tau^L, \tau^R)$ and $U^R(\tau^R, \tau^L)$ denote the representative country's per-period payoffs from the relationship with its left and right partner, respectively, when all countries select the same tariff pair (τ^L, τ^R).[14] The representative country's total per-period payoff is given by: $W = U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L)$. For trade policies (τ^L, τ^R) to be sustainable, they must provide each country with no incentive to defect; that is, the discounted welfare under these policies must be no less than the discounted welfare achieved by defecting and thereafter reverting to the punishment phase. Let us focus first on the incentive to defect against the left partner. If a country is to deviate against its left partner, it will deviate to its static optimum τ^{NL}. The gain from cheating is then given by $G^L(\tau^L, \tau^R) = U^L(\tau^{NL}, \tau^R) - U^L(\tau^L, \tau^R)$. Given that this deviation will be followed by a permanent reversion to the static Nash tariffs, the cost of future punishment is given by $\delta L^L(\tau^L, \tau^R)/(1-\delta)$, where $L^L(\tau^L, \tau^R) = [U^L(\tau^L, \tau^R) - U^L(\tau^{NL}, \tau^{NR})]$. Thus, the sustainability condition in the relationship with the left partner is:

$$(\mathrm{IC}^L)\quad G^L(\tau^L, \tau^R) \leq \delta L^L(\tau^L, \tau^R)/(1-\delta).$$

Similarly, the sustainability condition vis-à-vis the right partner can be expressed as:

$$(\mathrm{IC}^R)\quad G^R(\tau^R, \tau^L) \leq \delta L^R(\tau^R, \tau^L)/(1-\delta).$$

The most cooperative trade policies under BEM are the ones that maximize the payoff of the representative country, $W = U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L)$, subject to (IC^L), (IC^R), and the nonnegativity constraints $\tau^k \geq 0$ $(k = L, R)$.

Now consider a MEM, in which any defection is followed by a permanent reversion to the Nash tariffs in both relationships that involve the defecting country. Since any defection will be punished by both partners, if a country is to defect at all, it will defect against both partners simultaneously. Therefore, trade policies (τ^L, τ^R) can be supported by a MEM

[13] As we know from Dilip Abreu (1988), there may exist more severe (credible) punishments than permanent Nash reversions. Whether the analysis is robust to more general punishment strategies is an open question, but my conjecture is that considering more severe punishments would not change the qualitative results on the comparison between multilateral and bilateral enforcement.

[14] In terms of the notation introduced in Section I, we have: $U^L(\tau^L, \tau^R) \equiv U_m(\tau^L; z^R) + U_x(\tau^R; z^L)$ and $U^R(\tau^R, \tau^L) \equiv U_m(\tau^R; z^L) + U_x(\tau^L; z^R)$.

if and only if they satisfy the following incentive constraint:

$$(\mathrm{IC}^{\mathrm{M}})\ G^L(\tau^L, \tau^R) + G^R(\tau^R, \tau^L) \leq \delta[L^L(\tau^L, \tau^R) + L^R(\tau^R, \tau^L)]/(1-\delta).$$

The most cooperative trade policies under MEM are the ones that maximize $W = U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L)$ subject to $(\mathrm{IC}^{\mathrm{M}})$ and the nonnegativity constraints $\tau^k \geq 0$ $(k = L, R)$. I assume that δ is not too close to one, so that the nonnegativity constraints are not binding, under either BEM or MEM.[15] The next proposition (proved in the Appendix) summarizes the interesting findings.

PROPOSITION 1: *(i) In the presence of bilateral imbalances of power* $(z^L \neq z^R)$, *countries can sustain a higher symmetric equilibrium payoff with multilateral enforcement than with bilateral enforcement.*

(ii) Under bilateral enforcement, the weaker partner makes a larger "concession" than the stronger partner $(\tau^{NL} - \tau^L > \tau^{NR} - \tau^R)$. *The reverse is true under multilateral enforcement* $(\tau^{NL} - \tau^L < \tau^{NR} - \tau^R)$.

(iii) Absent power imbalances $(z^L = z^R)$, *bilateral and multilateral enforcement are equally efficient.*

Proposition 1 is illustrated in Figure 2. The curves IC^{L} and IC^{R} represent the two bilateral incentive constraints under BEM, taken with equality: when trade policies lie on IC^{L} (respectively IC^{R}), the representative country is indifferent between defecting and cooperating with its left (respectively right) partner; the area between curves IC^{L} and IC^{R} represents the set of sustainable outcomes under BEM. Given that the nonnegativity constraints are not binding, both IC^{L} and IC^{R} are binding under BEM, hence the optimal BEM agreement is given by point B. Point B lies below the 45-degree line that traces through the Nash point (line NE), which means that the weaker partner makes a larger concession than the stronger partner $(\tau^{NL} - \tau^L > \tau^{NR} - \tau^R)$. Under MEM, on the other hand, the only relevant incentive constraint ensures that a country has no temptation to defect against its left *and* right partners simultaneously. The curve IC^{M} represents this incentive constraint taken with equality; the sustainable set is given by the area northeast of curve IC^{M}. Drawn in Figure 2 are also some iso-W curves, i.e., curves along which W is constant. The optimal MEM agreement is identified by the point of tangency between the IC^{M} curve and the iso-W curve closest to the origin (point M). Point M lies on the line connecting the origin with the Nash point, which implies that under MEM the stronger partner makes a larger concession than the weaker partner $(\tau^{NL} - \tau^L < \tau^{NR} - \tau^R)$. Moreover, point M lies on a lower iso-*W* curve than point B, hence multilateral enforcement can sustain a higher symmetric payoff than bilateral enforcement.

To gain intuition, it is useful to start from the benchmark case of balanced power [Proposition 1(iii)].[16] Multilateral enforcement implies that any deviation will be met with punishment by both partners. Anticipating this, if a country is to deviate at all, it might as well deviate against both partners. Hence, multilateral enforcement doubles the loss from defecting relative to bilateral enforcement, but the gain from defecting is also doubled, thus the two enforcement mechanisms are equally effective. Proposition 1 (iii) makes clear that the gains from multilateral enforcement do not arise simply from the fact that defectors are punished more severely than under bilateral

[15] If δ is close to one, free trade can be sustained both with BEM and with MEM, and the comparison is uninteresting. There is also an interval of δ for which free trade can be sustained by MEM and not by BEM, but I am ignoring it for simplicity.

[16] Bendor and Mookherjee (1990) have already proved, although in a different application, the result that third-party sanctions provide no gains in a fully symmetric and separable game. Bendor and Mookherjee also look at an asymmetric game in which third-party sanctions are beneficial, but this has a very different structure from the power-imbalances game analyzed here. Also, Proposition 1 (iii) has a similar flavor as B. Douglas Bernheim and Michael Whinston's (1990) result that two firms attempting to collude on several markets have nothing to gain from cross-market punishment threats if markets are separable and firms are symmetric.

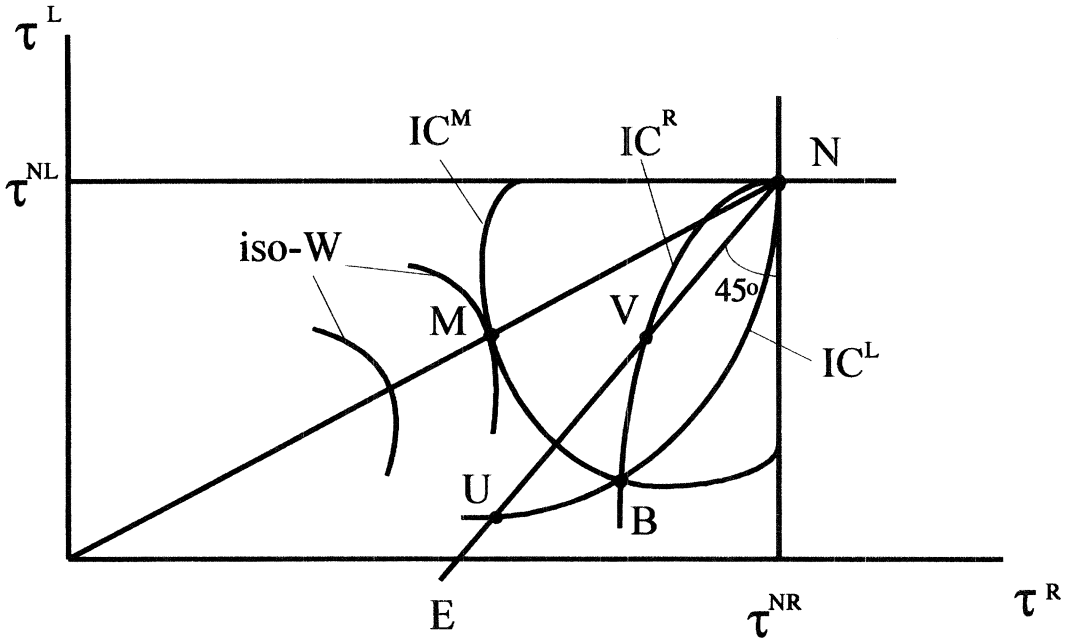

FIGURE 2. COMPARISON BETWEEN MULTILATERAL ENFORCEMENT AND BILATERAL ENFORCEMENT

enforcement, but arise specifically from the presence of power imbalances. If power imbalances are small, the gains from multilateral enforcement are also small.[17]

To understand why multilateral enforcement is beneficial in the presence of imbalanced power, it is useful to consider the following extreme case. Suppose the endowment structure is such that trade is perfectly triangular: a country imports only from its right partner and exports only to its left partner. In this setting, a country stands to lose a great deal from a trade war with its left partner, but gains from a trade war with its right partner. Here a BEM cannot sustain any trade liberalization, because in each bilateral relationship the "weak" partner has no capacity to punish defections from the "strong" partner. On the other hand, in a MEM each country acts as third-party enforcer in the relationship in which it is not directly involved, and some cooperation can be sustained.

I can now return to the question posed at the beginning of this section: Is there scope for an agency that verifies and publicizes violations of the agreements? Proposition 1 suggests that information dissemination can facilitate cooperation to the extent that there are imbalances of power in the trading system. In this regard, one could object that in reality tariffs are quite transparent, and changes in tariffs are readily observable. However, the model is best interpreted in a broader way. First, many nontariff import barriers are fairly opaque, especially to third countries. Second, trade agreements often allow countries to increase import barriers under certain contingencies; even if the import barrier per se is transparent, it may be hard for third countries to ascertain the exact

[17] It may be interesting to note that, if the model is extended to allow for monitoring costs, and multilateral monitoring is more costly than bilateral monitoring, then the net gains from a multilateral enforcement system are positive only if power imbalances are large enough.

underlying circumstances, hence whether or not the increase in protection is consistent with the agreement.

A. *International Transfers*

The reason why bilateral enforcement is less effective than multilateral enforcement is that, in each bilateral relationship, the gains from cooperation accrue unevenly to the two partners. Thus, a conjecture might be that, if pairs of countries could split evenly the gains from cooperation using international transfers (with the country that gains less from cooperation receiving the payment), then power imbalances would be neutralized, and multilateral sanctions would become superfluous. This conjecture can be examined within the model, if the space of feasible agreements is enlarged to include the possibility of (nonbinding) transfers. To make the relevant point it suffices to look at the case of triangular trade, where each country imports only from its right partner and exports only to its left partner. In this setting, the gains from cooperation can be redistributed at the bilateral level if each country makes a transfer to its left partner. I will focus on symmetric and stationary agreements of the following form: in each period, each country charges an import tariff τ on the good imported from its right partner, and simultaneously makes a payment Y to its left partner.[18] Assume again that δ is not too close to one, so that free trade cannot be sustained by MEM. The question is: if τ^M is the lowest tariff sustainable by MEM, does there exist a transfer Y^* such that the agreement (Y^*, τ^M) can be sustained by BEM?

To answer this question, let us first derive the lowest tariff sustainable by MEM. If all countries select tariff τ, the representative country's payoff is given by $U_m(\tau) + U_x(\tau)$, where the first (respectively second) term represents the per-period value of the relationship with the right (respectively left) partner. Here, there is only one deviation to consider, which yields a one-time gain of $U_m(\tau^N) - U_m(\tau)$. The per-period loss from defecting is given by $U_x(\tau) + U_m(\tau) - U_x(\tau^N) - U_m(\tau^N)$. After some algebra, the incentive constraint can be written as:

$$(\mathrm{IC}^{\mathrm{M}}) \qquad G_m(\tau) \leq \delta L_x(\tau),$$

where $G_m(\tau) \equiv U_m(\tau^N) - U_m(\tau)$ and $L_x(\tau) \equiv U_x(\tau) - U_x(\tau^N)$. Given the assumption that δ is sufficiently small, so that free trade cannot be sustained by MEM, it is direct to verify that the most cooperative tariff τ^M is the one that satisfies $(\mathrm{IC}^{\mathrm{M}})$ with equality.

Next, consider a BEM-cum-transfers agreement, (τ, Y). To write the two bilateral incentive constraints, notice that: (i) if a country defects against its right partner, it gains $G_m(\tau)$ today, and loses $(Y - G_m(\tau))$ in each period in the future;[19] (ii) if a country defects against its left partner (i.e., it witholds the transfer) it saves the transfer Y today and loses $(L_x(\tau) - Y)$ in each period in the future. The two incentive constraints can then be written as:

$$(\mathrm{IC}_{\mathrm{R}}^{\mathrm{B/Y}}) \quad G_m(\tau) \leq \delta[Y - G_m(\tau)] / (1 - \delta)$$

$$(\mathrm{IC}_{\mathrm{L}}^{\mathrm{B/Y}}) \quad Y \leq \delta[L_x(\tau) - Y] / (1 - \delta).$$

Now, suppose a transfer Y^* exists such that (Y^*, τ^M) satisfies these two constraints. From $(\mathrm{IC}_{\mathrm{L}}^{\mathrm{B/Y}})$, it must be $Y^* \leq \delta L_x(\tau^M)$. Since τ^M satisfies $G_m(\tau^M) = \delta L_x(\tau^M)$, it must be that $Y^* \leq G_m(\tau^M)$. But then the right-hand side of $(\mathrm{IC}_{\mathrm{R}}^{\mathrm{B/Y}})$ is negative, and hence this constraint is violated. This establishes that transfers cannot fully substitute for multilateral enforcement.

Next I examine whether transfers can at all facilitate cooperation under bilateral enforcement. The two bilateral incentive constraints are clearly binding at the most cooperative tariff; imposing equality in both of them and eliminating Y from the system, one finds that the lowest tariff sustainable by BEM-cum-

[18] The qualitative result of this section would not be affected if tariffs and transfers were chosen sequentially, rather than simultaneously.

[19] To see this, notice that the static Nash equilibrium is given by $\{\tau = \tau^N, Y = 0\}$, thus the per-period loss in the relationship with the right partner is $Y + U_m(\tau) - U_m(\tau^N) = Y - G_m(\tau)$.

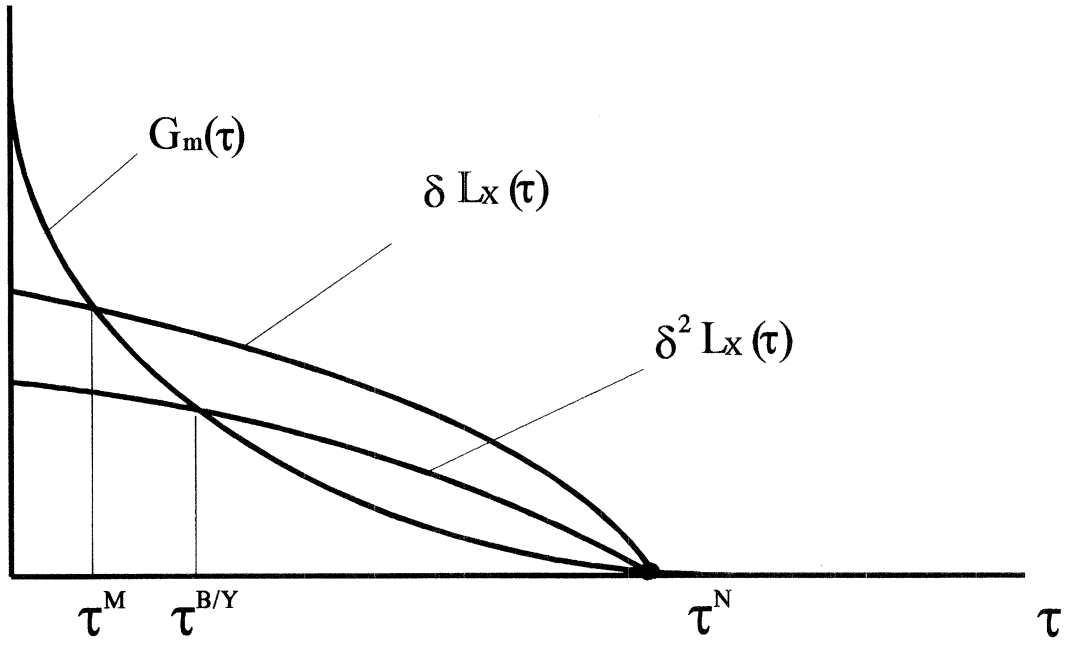

FIGURE 3. COMPARISON BETWEEN MULTILATERAL ENFORCEMENT AND BILATERAL ENFORCEMENT-CUM-TRANSFERS

transfers, $\tau^{B/Y}$, is determined by the following equality:

$$(IC^{B/Y}) \qquad G_m(\tau) = \delta^2 L_x(\tau).$$

Comparing $(IC^{B/Y})$ and (IC^{M}) graphically (see Figure 3), it is easy to verify that $\tau^{B/Y}$ lies between τ^{M} and τ^{N}. It is interesting to note that the transfer required to sustain $\tau^{B/Y}$ is $Y^{B/Y} = \delta L_x(\tau^{B/Y})$: the optimal transfer is not the one that splits the bilateral gains from liberalization, i.e., $Y = (G_m + L_x)/2$, as intuition might suggest. The reason is that, while this transfer equalizes the bilateral gains from cooperation, it does not equalize the gains from bilateral defection. In fact, there is *no* transfer that equalizes both the losses and the gains from bilateral defection. Thus, if δ is relatively small, the optimal transfer has to trade off between the two margins, and neither of them is equalized at the optimum. The next proposition summarizes the main finding.

PROPOSITION 2: *International transfers are an imperfect substitute for multilateral enforcement.*

Proposition 2 states that the use of transfers can mitigate, but not neutralize, the effects of bilateral imbalances of power. Intuitively, the reason why transfers cannot completely neutralize bilateral imbalances of power is that transfers, like trade policies, have to be self-enforcing. This imposes a constraint on the size of the transfer, since countries would be tempted to withhold transfers that are too large. The analysis shows that such a constraint is binding, that is, the transfer that neutralizes the imbalance of power is too high to be self-enforcing.

B. *Optimal Third-Party Sanctions*

In this section I take a closer look at the role of third-party sanctions. I will argue that, for the multilateral enforcement system to work at its best, third-party sanctions need only be of limited severity, and threatened only for certain violations.

Consider the basic model without international transfers. Again, I focus on equilibria in which all countries charge the same tariff pair (τ^{L}, τ^{R}) in all periods along the equilibrium

path, and punishments take the form of reversions to the static Nash tariffs. A simple way to parametrize the severity of bilateral and third-party punishments is by their duration; I will refer to a permanent punishment as "full," and to a temporary punishment as "partial." Let T_B^L denote the duration of the sanctions inflicted by the defector's left partner as bilateral punishment, and T_T^L the duration of the sanctions inflicted by the defector's left partner as third-party punishment; analogous notation applies for the right partner. Also, let $\underline{\tau}^M \equiv (\tau_L^M, \tau_R^M)$ denote the most cooperative tariffs under MEM. The following result (proved in the Appendix) establishes that full third-party punishment is not needed to support $\underline{\tau}^M$.

PROPOSITION 3: *The optimal tariff pair, $\underline{\tau}^M$, can be sustained by a combination of full bilateral punishment and partial third-party punishment.*

This result can be explained heuristically as follows. Under multilateral enforcement, there are three relevant incentive constraints for the representative country: two bilateral constraints, which require that the country have no temptation to defect against the left partner *or* the right partner; and a multilateral constraint, which requires that it have no temptation to defect simultaneously against the two partners. These constraints can be written as:

$$(\mathrm{IC}^{\mathrm{L}}) \qquad G^L \leq \Lambda_B^L + \Lambda_T^R$$

$$(\mathrm{IC}^{\mathrm{R}}) \qquad G^R \leq \Lambda_B^R + \Lambda_T^L$$

$$(\mathrm{IC}^{\mathrm{M}}) \qquad G^L + G^R \leq \Lambda_B^L + \Lambda_B^R,$$

where G^L denotes the one-time gain from defecting against the left partner, Λ_B^L the *discounted* loss inflicted by the left partner as bilateral punishment, and Λ_T^L the discounted loss inflicted by the left partner as third-party punishment (these loss terms are increasing in the duration of the sanctions); analogous notation applies for the right partner. First suppose that both bilateral and third-party sanctions are maximal: then, only (IC^{M}) will be binding, because a government would rather defect against both partners than against a single partner. Now note that the severity of third-party sanctions does not affect (IC^{M}), because in case of simultaneous deviation both injured countries punish to the full extent. Therefore, third-party sanctions can be reduced below the maximal level (while keeping bilateral sanctions at the maximal level) without affecting the sustainability of $\underline{\tau}^M$.

What is the minimum third-party punishment sufficient to sustain $\underline{\tau}^M$? To answer this question, first notice that Λ_T^L and Λ_T^R each can be lowered until the corresponding bilateral constraint begins to bind, or in other words, until a country is indifferent between defecting against a single partner and defecting against both partners simultaneously. It is not hard to show that Λ_T^R can be lowered to zero; in words, third-party punishment is not necessary for violations by a weak partner against a strong partner. On the other hand, the minimum sufficient level of Λ_T^L (i.e., the severity of third-party punishment for violations of the strong against the weak) is positive, and can be shown to be increasing in the size of the power imbalance: if power is nearly balanced, Λ_T^L need only be small.[20]

The result can be illustrated with the aid of Figure 4, which builds on Figure 2. Under full third-party punishment, we know that the optimum is given by point M, the point of tangency between the curve IC^{M} and the iso-W curve closest to the origin. With no third-party punishments ($\Lambda_T^L = \Lambda_T^R = 0$), the three incentive constraints (IC^{M}, IC^{L}, and IC^{R}) cross at point B, that is the optimal tariff pair under BEM. As Λ_T^L is increased above zero, the IC^{R} curve shifts up. There is a critical level of Λ_T^L, say Λ_T^{L*}, such that IC^{R} traces through point M (curve $\mathrm{IC}^{\mathrm{R}*}$); this means that the tariff pair M can be supported with third-party

[20] It is worth noting that, if there are many countries in the trading system, the amount of punishment that needs to be threatened *by each third country* will be small. More precisely, if one considers an extension of the model in which the representative country has n "weaker" partners and n "stronger" partners, one can show that the minimum sufficient amount of punishment that needs to be threatened by each third country becomes negligible as n gets large.

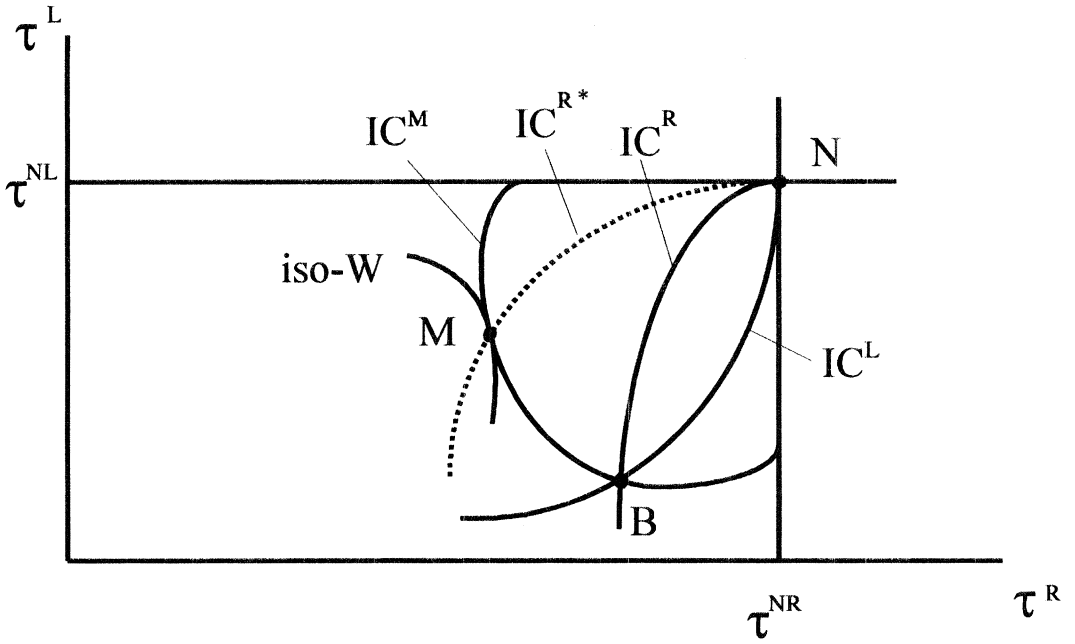

FIGURE 4. ILLUSTRATION OF PROPOSITION 3

punishments $(\Lambda_T^{L*}, 0)$. Increasing Λ_T^L any further, or making Λ_T^R positive, has no benefit.

I define third-party sanctions as *minimal* if they are the least severe that, in combination with full bilateral sanctions, can sustain $\underline{\tau}^M$. In terms of the notation just introduced, these are given by $(\Lambda_T^{L*}, 0)$. I will argue that minimal third-party sanctions become strictly optimal if two slight changes are made to the model, namely: (i) monitoring of trade policies is slightly imperfect, and (ii) punishing is slightly more costly for a third party than for a second party.

To formalize the monitoring imperfection in the simplest way possible, suppose that, if a country selects tariff τ^k $(k = L, R)$, the remaining countries perceive the true level τ^k with probability $1 - q$, but with a small probability q they mistakenly perceive the noncooperative level τ^{Nk}, regardless of the true policy. Next suppose that, for unmodeled reasons, a third country that participates in the punishment process incurs a small extra loss ϵ that the defector and the second party do not incur. More precisely, suppose that the per-period bilateral payoff of a third country involved in the punishment of its left (respectively right) partner is given by $U^L(\tau^{NL}, \tau^{NR}) - \epsilon$ [respectively $U^R(\tau^{NR}, \tau^{NL}) - \epsilon$]; the parameter ϵ does not affect the payoffs of the defector or the injured country. One possible interpretation of ϵ is that implementing trade sanctions is politically more costly for a third-country government than for the government of the injured country. This assumption is ad hoc, but it seems a reasonable one, given that its only role is that of a tiebreaker. This small extra cost of third-party sanctions will translate into a large difference between the optimal bilateral and third-party sanctions. In Corollary 1 (proved in the Appendix), I refer to an "optimal" enforcement mechanism as one that sustains the highest symmetric payoff within the class of punishment strategies I am focusing on.

COROLLARY 1: *In the limit, as q and ϵ approach zero, the optimal enforcement mechanism entails minimal third-party sanctions.*

Corollary 1 states that, if monitoring is slightly imperfect and if punishing is slightly more costly for third parties than for second parties, third-party sanctions should be kept at their "minimal" levels. The intuition for this result is simple. Proposition 3 suggests that, with $q = \epsilon = 0$, there is a continuum of equivalent enforcement mechanisms, along two dimensions. One is the extent of the total (bilateral plus third-party) punishment: above a certain critical level, the severity of total punishment is irrelevant. The other dimension is the composition of total punishment between the bilateral and third-party components. The monitoring imperfection breaks the indifference in favor of mechanisms that minimize total punishment; this is because undue punishments are occasionally triggered along the equilibrium path, and hence overpunishment is costly. The small extra cost of third-party sanctions breaks the indifference with respect to the composition of total punishment, in favor of mechanisms that minimize the third-party component.[21]

The results of this section suggest that punishment norms in the trading system should be "parsimonious": third-party sanctions should be *minimal* in terms of severity, and *selective*, in the sense that they should be threatened only for violations that are harder to deter with bilateral sanctions alone, such as violations of the strong against the weak.

In the model, third-party sanctions take the simple form of temporary reversions to higher tariffs. In a broader interpretation of the model, one can think of third-party sanctions as possibly taking different forms. For example, after a violation of the agreement, third countries might be more reluctant to enter new agreements with the offending country or, if the current agreement entails gradual trade liberalization, they might slow down the pace of the ongoing liberalization process.

Before moving to the next section, I should briefly discuss a possible criticism of the analysis, namely, that it does not allow for renegotiation of the agreement after a violation.[22] Imposing a renegotiation-proof constraint would lower the maximum sustainable level of cooperation, but is unlikely to upset the qualitative results of the model. In fact, the gains from multilateral enforcement may become more important in the presence of renegotiation concerns, for the following reason. We know from the previous analysis that the most cooperative agreement entails a combination of maximum bilateral sanctions and partial third-party sanctions; since imposing a renegotiation-proof constraint amounts to imposing a cap on the severity of the punishment that can be credibly threatened by each partner, it is likely to result in a reduction of bilateral sanctions and an increase of third-party sanctions. The optimal amount of third-party sanctions relative to bilateral sanctions would then increase as renegotiation considerations are introduced. A second remark is that a renegotiation-proof constraint is likely to constitute less of a problem when power imbalances are deeper, hence when third-party sanctions are more important. The reason is that, with deep power imbalances, punishments are very asymmetric: the punishing country gets hurt much less than the punished country, hence the former has a limited incentive to renegotiate.

C. *Imperfect Commitment to the Verification Mechanism*

Suppose that an independent verification agency like the DSP is available, but the verification activity cannot be directly enforced because a government is ultimately free to foreclose the agency's access to the relevant information. Would this undermine the multilateral enforcement mechanism? The issue can be phrased differently: For multilateral enforcement to be effective, countries must com-

[21] Note that this finding is very different from Edward Green and Robert Porter's (1984) result that optimal punishments are less than maximal in the presence of imperfect monitoring. In Green and Porter's (1984) model, the optimal punishment is close to maximal when uncertainty is small, whereas here it is close to "minimal" when uncertainty is small.

[22] Several notions of renegotiation have been proposed in the literature (see in particular Bernheim and Debraj Ray, 1989; Joseph Farrell and Eric Maskin, 1989). See Rodney Ludema (1990) for an examination of the consequences of renegotiation in the context of the GATT's dispute settlement procedure.

ply not only with *substantive* rules (trade policies), but also with *procedural* rules, such as facilitating the verification activity. Is the effectiveness of multilateral enforcement diminished by the fact that procedural rules must be self-enforcing?

In this section I will argue that this nonenforceability problem need not impose a binding constraint on the multilateral system. Consider the following modification of the basic model. Suppose a government's violation is automatically verified and publicized to the trading community by a verification agency, unless the government blocks the transmission of information. In every period, each government chooses its trade policies and (simultaneously) chooses whether to block the transmission of information (the result does not change if the blocking option is exercised following the trade-policy decision). Exercising the blocking option is costless. If the blocking option is exercised, the verification agency can publicize this procedural violation before the end of the period. The following result (proved in the Appendix) is obtained.

PROPOSITION 4: *If governments have the option to block the transmission of information, the most cooperative tariff pair is* $\underline{\tau}^M$, *the same as in the absence of the blocking option.*

This result relies on the very simple nature of the procedural rule considered here. However, the analysis suggests a conclusion that ought to be valid quite generally: when international cooperation must take place on "substantive" as well as "procedural" dimensions, the self-enforcement constraint on procedural rules need not be binding, or in other words, procedural rules need not "compete" with substantive rules for the scarce supply of enforcement power, as long as (i) the DSP can publicize procedural violations, and (ii) a government does not benefit from a procedural violation *in itself*, but opts to violate procedures only to avoid sanctions for substantive violations.

One can certainly think of situations in which these conditions are not satisfied. For example, if the cost of operating the institution is significant, countries will be required to provide financial contributions. Contributing to the institution's budget can be thought of as a procedural rule, for which condition (ii) is not satisfied. In this case, the self-enforcing constraint on the procedural rule may well be binding, and reduce the sustainable level of cooperation on the substantive dimension. Also, one can think of procedural violations that are not publicly observable, so that condition (i) is not satisfied. For example, if a government is required to provide all the relevant documentation to the DSP, a government may be able to withold some of this information without being detected.

III. Efficiency Gains from Multilateral Rule-making

In this section I focus on the implications of multilateralism at the level of trade negotiations. Trade negotiations can be thought of as a process of bargaining over the set of self-enforcing agreements, in the sense that governments select a tariff configuration subject to the constraint that no participant has an incentive to defect from the agreed-upon tariffs. This conceptualization is useful to disentangle the benefits of multilateralism at the rule-making level from those at the enforcement level. Assume governments engage in Nash bargaining, with a threat point given by the static Nash equilibrium payoffs. *Multilateral bargaining* is defined as a single Nash bargaining game that involves all countries. *Bilateral bargaining* is defined as a web of (three) simultaneous bilateral Nash bargaining games, where each pair bargains taking all other tariffs as given.[23] I assume that bargaining powers are equal both at the multilateral level and at the bilateral level.[24] In order to illustrate the importance of multilateral negotiations, I will compare a fully multilateral system with a system in which enforcement is multilateral but negotiations are bilateral.

[23] Since the trading system is separable, this reduces to applying the Nash solution separately to each pair of countries.

[24] It might be more realistic to assume that, at the bilateral level, the "stronger" country has more bargaining power. However, this would only strengthen (one can show) the main result of this section.

To sharpen intuition, consider the basic model of Section I and suppose that δ is close to one, so that free trade is sustainable. Suppose further that bilateral trade is very unbalanced, so that one country would "win" the bilateral trade war (i.e., would be better off under a trade war than under reciprocal free trade). Multilateral bargaining would select free trade, which is the symmetric Pareto-optimal outcome. On the other hand, bilateral bargaining would *not* select free trade, because free trade does not Pareto improve over the disagreement payoffs at the bilateral level. Thus, in this example, multilateral bargaining yields efficiency gains relative to a bilateral bargaining procedure. Bilateral bargaining generates a bias in favor of the "strong" country in each bilateral relationship, which translates into an inefficiency at the global level.

When δ is relatively low, so that enforcement problems come into play, multilateralism at *both* the bargaining level and the enforcement level becomes essential to achieve efficient outcomes. To make this point clear, suppose that neither $\tau^L = 0$ or $\tau^R = 0$ can be sustained. With reference to Figure 5, the set of sustainable tariffs is the area that lies northeast of curve IC^M. Focusing on multilateral bargaining first, the disagreement payoff is $(U^{NL} + U^{NR})$ for all governments, where $U^{NL} \equiv U^L(\tau^{NL}, \tau^{NR})$ and $U^{NR} \equiv U^R(\tau^{NR}, \tau^{NL})$. Hence the Nash bargaining solution selects the point in the sustainable set that maximizes the Nash criterion $[(U^L + U^R) - (U^{NL} + U^{NR})]^3$, or equivalently, the one that maximizes $(U^L + U^R)$ (point MB). Under bilateral bargaining, on the other hand, each pair of countries picks the tariff pair in the sustainable set that maximizes the Nash criterion $(U^L - U^{NL})(U^R - U^{NR})$ (point BB). If $z^L \neq z^R$, point BB is different from point MB, and clearly entails a lower total payoff for the representative country. The next proposition follows readily.

PROPOSITION 5: *Multilateral bargaining yields a Pareto-superior outcome relative to bilateral bargaining if and only if* $z^L \neq z^R$.

Proposition 5 identifies a failure of decentralized bargaining procedures. The interesting aspect of this result is that bilateral bargaining is locally efficient (i.e., it is efficient conditional on the outcome of other bilateral bargains) and there are no externalities across bilateral relationships, yet it is globally inefficient. The failure of decentralized bargaining is due solely to the presence of bilateral imbalances. Of course, in reality there are important nonseparabilities in the trading system, generated by the trade-diverting effects of trade policies, but the presence of nonseparabilities is likely to strengthen the result of Proposition 5, as I will argue in Section IV.[25] As already explained in the introduction, the inefficiency of bilateral bargaining highlighted by the model is akin to the inefficiency caused by incomplete markets.[26]

A natural question concerns the role of international transfers, which have not been considered in this section. The result of Proposition 5 relies on the fact that governments' utility is not transferable. One can show that if *enforceable* transfers are available, bilateral bargaining produces efficient outcomes. The reason is that in this case utility becomes transferable, and hence efficiency concerns become separate from distribution concerns. Under bilateral bargaining, each pair

[25] In the literature on bargaining there are several models—some adopting the Nash bargaining approach, others adopting the noncooperative approach—that point out inefficiencies of decentralized bargaining mechanisms (see Martin Osborne and Ariel Rubinstein, 1990, for a comprehensive survey), but these inefficiencies are always driven by externalities across bilateral relationships, whereas in the present model there are no such externalities.

[26] It is worth relating the result of this section to Andrew Caplin and Kala Krishna's (1988) model of bilateral bargaining with most-favored-nation (MFN) extension. They formalize the well-known problem that the MFN requirement creates a free-rider problem, since a bilateral agreement between countries A and B generates a "free" tariff reduction for country C. They find that bilateral bargaining-cum-MFN yields less efficient outcomes than unconstrained bilateral bargaining. It should be noticed that, in their model, unrestricted bilateral bargaining yields the same outcome as multilateral bargaining, hence no failure of decentralized bargaining *as such* is identified. Their result is complementary with the one presented here; taken together, they suggest that: (i) multilateral bargaining may be necessary to achieve efficient outcomes, and (ii) multilateralism should not be introduced in the form of bilateral bargaining with MFN extension; this is likely to make things worse rather than better.

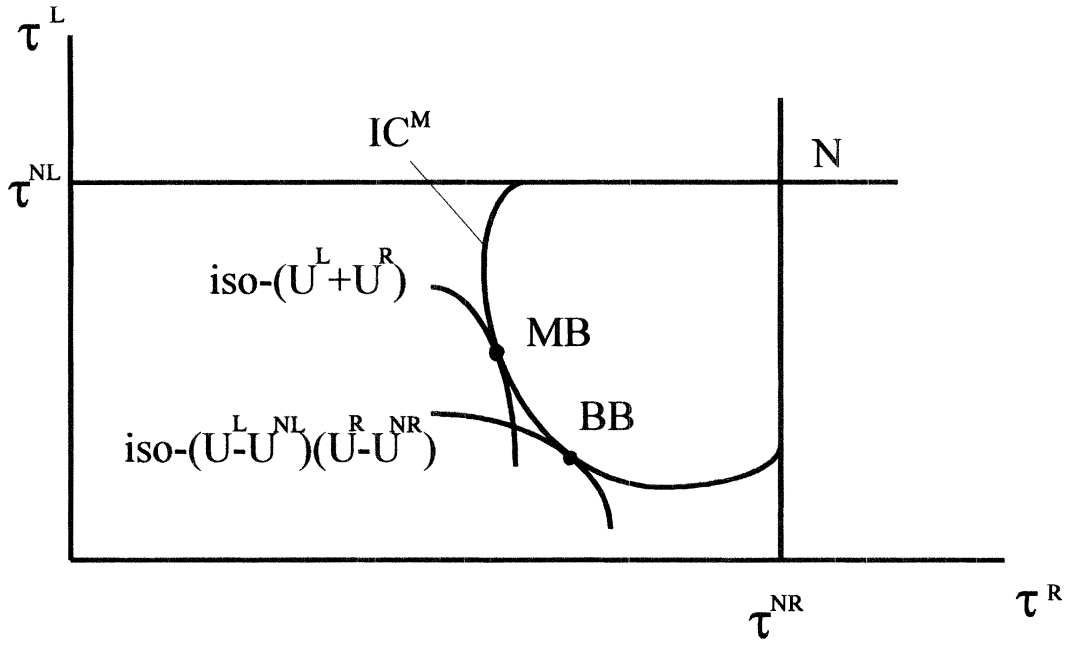

FIGURE 5. COMPARISON BETWEEN MULTILATERAL BARGAINING AND BILATERAL BARGAINING

of countries selects the tariff pair that maximizes the joint surplus ($U^L + U^R$), and transfers are used to redistribute the surplus, thus the outcome is globally efficient. However, if transfers are not enforceable and δ is relatively low, things are different: using a similar argument as in Section II, subsection A, one can show that transfers can only mitigate, but not remove, the inefficiency caused by bilateral bargaining. Overall, Proposition 5 should be interpreted as applying to situations where: (i) utility is not transferable, or (ii) utility is transferable, but transfers are not enforceable, and δ is relatively low.

This analysis can be useful in interpreting the current debate about multilateralism in trade negotiations. Some of the informal literature, for example Jagdish Bhagwati (1990) and L. Alan Winters (1990) suggests that a multilateral approach may have significant distributional effects, as it favors countries that have weak bargaining positions in a bilateral-negotiation setting. The analysis here indicates that there may also be an important *efficiency* rationale for a multilateral approach, in addition to any possible distributional motivation.

IV. Extensions

In this section, I will discuss how the results concerning the gains from multilateral enforcement and rule-making generalize to richer settings. I will focus first on the gains from multilateral enforcement, discussing the implications of: (i) global imbalances of power, (ii) trading blocks, and (iii) trade diversion. Then, at point (iv), I will address the gains from multilateral rule-making.

(i) The model focused on local imbalances of power, in the context of a globally symmetric system. One can construct a version of the model that allows for *global* imbalances of power. Suppose that at each point in time one country is stronger than all others, but the identity of the "leader" is subject to change over time. In each period, there is a probability that the current leader will be replaced by

another country. Also suppose that, at the time the agreement is struck, each country has the same probability of becoming the leader. In this setting, the current leader is a natural enforcer of the agreements, as it possesses abundant third-party enforcement power.

In my working paper (Maggi, 1996) I show that, in this setting, a multilateral enforcement mechanism allows countries to achieve a higher symmetric expected payoff than a bilateral enforcement mechanism. Each country provides third-party enforcement services when it finds itself in the position of leader, and receives them when it is in "weak" position. I also show that the benefits of multilateral enforcement are larger when the leadership is less stable; intuitively, if the leadership is very stable there is little scope for intertemporal exchange of enforcement power.

(ii) In the context of self-enforcing trade agreements, a *trading bloc* can be thought of as a group of countries that places an especially high value on future cooperation among themselves, relative to the short-run gain from erecting trade barriers, so that free trade can be sustained within the bloc. A situation in which multiple trading blocs can arise, for example, is one in which the volume of intragroup trade is expected to grow at a faster rate than that of intergroup trade; this is because free trade is easier to sustain when trade is expected to grow faster, as the prospect of trade expansion makes the consequences of future punishment more important relative to the one-time gain from defecting. In the working paper I show that, when free trade is sustainable within trading blocs but not across blocs, multilateral enforcement is more effective than bilateral enforcement. The reason is that multilateral enforcement allows a transfer of enforcement power from intragroup relationships, where it is abundant, to across-group relationships, where it is scarce. The key mechanism that makes multilateral enforcement more efficient is that defections against nonmember countries trigger sanctions from member countries. Since losing cooperation with member countries has particularly damaging consequences, this threat can help enforce deeper cooperation with nonmember countries.

One can also show that, if global free trade cannot be sustained, the optimal multilaterally enforced agreement does not entail free-trade areas, even though the optimal bilaterally enforced agreement does. To understand this, consider the optimal bilaterally enforced agreement, which entails free trade within blocs and positive tariffs across blocs. Since a small deviation from free trade has a second-order welfare cost, this agreement can be improved upon by reducing the interbloc tariff and introducing a small intrabloc tariff. This rearrangement of tariffs, however, can be sustained only through multilateral enforcement, since only in this way can countries redistribute enforcement power across bilateral relationships. This result suggests that, if enforcement power is a scarce resource, implementing free-trade areas implies an inefficiency in the use of this resource.

(iii) The gains from multilateral enforcement discussed thus far can be thought of as gains from *exchanging* enforcement power across bilateral relationships. There may also be gains from *aggregating* enforcement power in the trading system. Gains from aggregating enforcement power arise when collective punishments are proportionately more severe than bilateral punishments. This situation can arise when trade policies cause trade diversion.

One reason why multilateral punishment may be proportionately more severe than bilateral punishment can be illustrated by a simple example. Suppose there are only three countries, and that punishments take the form of trade embargoes. Under multilateral enforcement, a violation by country 1 against country 2 is punished by a double embargo against country 1, so that country 1 loses both export markets. Under bilateral enforcement, on the other hand, the same violation is punished with the interruption of trade only between countries 1 and 2; as a consequence, country 1's exports to country 3 will increase, partially substituting for the loss of market 2. For this reason, the loss from a bilateral embargo is less than half the loss from a multilateral embargo, and hence multilateral enforcement may allow countries to sustain lower trade barriers. The same effect is likely to arise also when punishments take the less extreme form of tariff increases.

Another reason why there may be gains from aggregating enforcement power is that import tariffs charged by different governments on the same good tend to be strategic complements, in the sense that a government's best-response tariff is increasing in the other governments' tariffs (this is pointed out also in Bagwell and Staiger, 1997). The reason is that, if country 1 raises its tariff on a good imported from country 2, country 1 will import a lower amount of that good, some trade will be diverted to country 3, and this will invite an increase of country 3's tariff on that good. Because of this strategic complementarity, the level of punitive tariffs tends to be higher if more governments participate in the punishment, thus multilateral punishment is proportionately more severe than bilateral punishment.

(iv) Turning to the gains from multilateral rule-making, Proposition 5 can be generalized in the following way. Consider a trading system with n countries, that is globally symmetric (in the sense that it looks identical from any one country's point of view) and has no trade diversion. In such a trading system, the presence of bilateral asymmetries is a necessary and sufficient condition for multilateral bargaining to provide efficiency gains relative to bilateral bargaining (Maggi, 1996). Interestingly, this means that gains from multilateral rule-making arise under a narrower set of conditions than gains from multilateral enforcement. For example, if the system is characterized by trading blocs [as in point (ii) above], but countries are pairwise symmetric, bilateral and multilateral bargaining are equally efficient; yet, there may be gains from multilateral enforcement, due to the presence of asymmetries across bilateral relationships.

If the trading system is nonseparable, so that tariffs have trade-diverting effects, the gains from multilateral rule-making relative to bilateral rule-making tend to be strengthened. This is intuitive, since in the presence of nonseparabilities there are externalities across bilateral relationships. In a bilateral negotiation, countries do not take these externalities into account, whereas a multilateral rule-making procedure internalizes them.

V. Conclusions

In this paper I have examined two potential benefits of a multilateral trade institution that lacks direct enforcement power: first, the institution can verify violations of the agreement and inform third parties; second, it can promote an efficient multilateral rule-making procedure. The model suggests that a multilateral approach to enforcement and rule-making is particularly important when there are strong power imbalances in bilateral trading relationships.

Since the model examines the *potential* benefits of a trade institution, a natural interpretation of the analysis is normative. Having said this, it is legitimate to ask: To what extent have the potential benefits been realized by the modern trading system?

There is some evidence that the GATT has effectively promoted a multilateralization of the bargaining process in recent decades. As Finlayson and Zacher (1981) write:

> Rule making (...) was dominantly bilateral in character during most of the GATT's first two decades ... However, the development of linear tariff negotiations and, even more so, the growing importance of NTBs in recent bargaining rounds have introduced a stronger component of multilateralism into decision making in the GATT. True, the major trading states continue to dominate the rule-making process, but more regime members now participate in any given negotiation.

A more controversial question is whether the activity of the GATT's dispute settlement procedure has improved the enforcement of trade agreements in the way that the model suggests. This question will not receive a full answer here, but a useful first step may be to take a look at some anecdotal evidence from the history of GATT disputes.

There are several examples of GATT disputes in which there was a clear imbalance of power between the disputants, and the DSP panel ruled in favor of the weaker country, inducing the stronger country to withdraw the GATT-illegal policy. I will first review

some of these cases, then discuss their interpretation.[27]

(1) In 1949, Chile filed a complaint against Australia for a subsidy on ammonium sulfate that allegedly violated the MFN rule. Australia complied. (2) In 1961, Brazil filed a complaint when the United Kingdom increased tariffs on bananas for non-Commonwealth countries. The United Kingdom was found in violation, and agreed to comply. (3) In 1972, Israel brought the United Kingdom before the DSP for GATT-illegal import restrictions on cotton textiles. The United Kingdom complied. (4) In 1979, Chile filed a complaint against the European Community (EC) for its illegal quotas on imports of apples from Chile. The EC did not renew the quota. (5) In 1980, India alleged that the United States had imposed a GATT-illegal countervailing duty without making the required determination of "material injury." The United States agreed to withdraw the duty before the DSP panel was formed. (6) In 1981, Hong Kong challenged the legality of EC quotas on a variety of Hong Kong's products, in particular quartz watches. The EC complied. (7) In 1993, Guatemala, Nicaragua, Costa Rica, Venezuela, and the Dominican Republic alleged that the EC had a GATT-illegal banana import regime, discriminating in favor of former EC colonies. The panel ruled against the EC, and this prompted a framework agreement on bananas, which essentially accommodated the Latin American countries' requests. (8) In 1995, Costa Rica filed a complaint against U.S. quota restrictions on textiles, introduced as "safeguard" measures, without making the required determination of "material injury." The United States withdrew the restriction a few months later.

How these cases should be interpreted is debatable. At a minimum, they indicate that the DSP had some impact on the behavior of governments. It is hard to imagine that the outcome would have been the same without the presence of the institution. A more difficult question is, just what role was played by the DSP? The interpretation I propose here is that the weak countries invoked the DSP to inform the whole trading community of the GATT-illegal policies, and the strong countries complied with the GATT panel for fear of a loss of reputation in the GATT arena.

However, one more step is needed to argue that this anecdotal evidence is consistent, at least in a broad sense, with the theory presented in this paper. One has to argue that the strong disputants could, at least in a future expected sense, benefit from the enforcement power of the weaker disputants. There are three reasons why this might be the case. First, a country that is economically weak today might become a significant trading partner tomorrow. For example, in the cases reviewed above, this could perhaps be claimed for countries like India and Brazil. Second, countries that dominate the trading scene today may decline tomorrow; after all, some decades ago the United Kingdom was the dominant country in the trading system, but is no longer. Third, even if a strong country does not expect to benefit from the enforcement power of a single small country, it may expect to benefit from the aggregation of small countries, when it comes to a dispute with a strong trading partner; for example, the United States might conceivably benefit from the enforcement pressures of the whole WTO community when it comes to disputes with the European Union or Japan.

APPENDIX

PROOF OF PROPOSITION 1:

I will prove statement (ii) first, showing that the most cooperative BEM tariff pair (point B in Figure 2) lies below the 45-degree line through N (line EN), while the most cooperative MEM tariff pair (point M) lies above it.

Let $U_m^L(\tau^L) \equiv U_m(\tau^L; z^R)$, $U_x^L(\tau^R) \equiv U_x(\tau^R; z^L)$, $U_m^R(\tau^R) \equiv U_m(\tau^R; z^L)$ and $U_x^R(\tau^L) \equiv U_x(\tau^L; z^R)$. Since these functions are all quadratic, they can be written as $U_j^k(x) = a_{jk}x^2 + b_{jk}x + c_{jk}$ $(j = m, x; k = L, R)$. If the demand function is given by $d(p) = \alpha - \beta p$, one can check that $a_{mL} = a_{mR} = -3\beta/8$, $a_{xL} = a_{xR} = \beta/8$, $b_{mL} = z^R/4$, $b_{mR} = z^L/4$, $b_{xL} = -z^L/4$, $b_{xR} = -z^R/4$ (the c_{jk} terms are not needed).

[27] The source of information for these cases is Hudec's (1993) book.

Consider first a BEM. The two bilateral incentive constraints can be written (after some algebra) as:

$$(\mathrm{IC}^{\mathrm{L}}) \quad U_m^L(\tau^{NL}) - U_m^L(\tau^L) \leq \delta[U_x^L(\tau^R) - U_x^L(\tau^{NR})]$$

$$(\mathrm{IC}^{\mathrm{R}}) \quad U_m^R(\tau^{NR}) - U_m^R(\tau^R) \leq \delta[U_x^R(\tau^L) - U_x^R(\tau^{NL})].$$

The most cooperative BEM tariffs maximize $W = U_m^L(\tau^L) + U_x^L(\tau^R) + U_m^R(\tau^R) + U_x^R(\tau^L)$ subject to $(\mathrm{IC}^{\mathrm{L}})$, $(\mathrm{IC}^{\mathrm{R}})$, and $\tau^k \geq 0$ $(k = L, R)$. Since the nonnegativity constraints are not binding (by assumption), constraints $(\mathrm{IC}^{\mathrm{L}})$ and $(\mathrm{IC}^{\mathrm{R}})$ are both binding. The most cooperative BEM tariffs must then satisfy these two constraints taken with equality. In Figure 2, the curves IC^{L} and IC^{R} represent the bilateral incentive constraints taken with equality. These curves intersect at two points: one is the Nash point N, and the other is point B, the most cooperative BEM tariff pair. To show that point B lies below line EN, it suffices to show that point V (the intersection between IC^{R} and EN) is closer to point N than point U (the intersection between IC^{L} and EN). To this end, I will argue that the value of τ^R at point V (say τ_V^R) is higher than at point U (say τ_U^R). To find τ_V^R, impose $\tau^{NL} - \tau^L = \tau^{NR} - \tau^R$ in the equation that defines IC^{R} and solve for τ^R, obtaining: $\tau_V^R = [(3 - \delta)z^L - 4\delta z^R]/3\beta(3 - \delta)$. With analogous procedure, one finds: $\tau_U^R = (3 - 5\delta)z^L/3\beta(3 - \delta)$. It is immediate to check that $\tau_V^R > \tau_U^R$, using the fact that $z^L > z^R$.

Next focus on a MEM. The incentive constraint $(\mathrm{IC}^{\mathrm{M}})$ can be written as:

$$(\mathrm{IC}^{\mathrm{M}}) \quad U_m^L(\tau^{NL}) - U_m^L(\tau^L) + U_m^R(\tau^{NR}) - U_m^R(\tau^R) \leq \delta[U_x^L(\tau^R) + U_x^R(\tau^L) - U_x^L(\tau^{NR}) - U_x^R(\tau^{NL})].$$

The most cooperative tariff pair is the one that maximizes $W = U_m^L(\tau^L) + U_x^L(\tau^R) + U_m^R(\tau^R) + U_x^R(\tau^L)$ subject to $(\mathrm{IC}^{\mathrm{M}})$ and $\tau^k \geq 0$ $(k = L, R)$. Given that the nonnegativity constraints are not binding, it is easy to show that $(\mathrm{IC}^{\mathrm{M}})$ is binding. The first-order conditions for this constrained maximization imply the following equality:

$$U_m^{L'}(\tau^L)U_x^{L'}(\tau^R) = U_m^{R'}(\tau^R)U_x^{R'}(\tau^L).$$

Graphically, this condition is that the IC^{M} curve is tangent to an iso-W curve. Plugging the quadratic specifications, this tangency condition reduces to: $\tau^L/\tau^R = z^R/z^L$, which is the equation of the line connecting point N with the origin. Since the most cooperative MEM tariff pair (point M) must lie on this line, it follows that point M lies above line EN.

Statement (i) follows from the facts that: (a) MEM is weakly more efficient than BEM, and (b) the most cooperative MEM tariff pair yields a different total payoff than the most cooperative BEM tariff pair. Statement (iii) can be established by inspection of Figure 2, noting that if $z^R = z^L$ the picture is perfectly symmetric, and the most cooperative BEM tariff pair (point B) coincides with the most cooperative MEM tariff pair (point M).

PROOF OF PROPOSITION 3:

Let $\underline{T}_B \equiv (T_B^L, T_B^R)$ and $\underline{T}_T \equiv (T_T^L, T_T^R)$. Fixing $\underline{T}_B = (\infty, \infty)$, the three relevant incentive constraints given third-party sanction durations (T_T^L, T_T^R) can be written as:

$$(\mathrm{IC}^{\mathrm{M}}) \quad U^L(\tau^{NL}, \tau^R) + U^R(\tau^{NR}, \tau^L) - U^L(\tau^L, \tau^R) - U^R(\tau^R, \tau^L) \leq \frac{\delta}{1-\delta}[U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L) - U^{NL} - U^{NR}],$$

$$(\mathrm{IC}^{\mathrm{L}}) \quad U^L(\tau^{NL}, \tau^R) - U^L(\tau^L, \tau^R) \leq \frac{\delta}{1-\delta}[U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L) - U^{NL} - (1 - \delta^{T_T^R})U^{NR} - \delta^{T_T^R+1}U^R(\tau^R, \tau^L)],$$

$$(\mathrm{IC}^R) \quad U^R(\tau^{NR}, \tau^L) - U^R(\tau^R, \tau^L)$$

$$\leq \frac{\delta}{1-\delta}\left[U^L(\tau^L, \tau^R) + U^R(\tau^R, \tau^L) - U^{NR} - (1-\delta^{T_T^L})U^{NL} - \delta^{T_T^L+1}U^L(\tau^L, \tau^R)\right].$$

Tariffs $\underline{\tau}^M$ are by definition the ones that maximize $U^L + U^R$ subject to the three incentive constraints above when we impose $T_T^L = T_T^R = \infty$. We need to show that $\underline{\tau}^M$ can be sustained also with some finite T_T^L and T_T^R.

If $T_T^L = T_T^R = \infty$, only the constraint (IC^M) is binding. To see this, observe that the right-hand side is the same in all three constraints, and the left-hand side of (IC^M) is strictly higher than that of the other two constraints. It remains to argue that we can decrease both T_T^L and T_T^R below infinity without violating any constraint and without decreasing the objective. But this is immediate, because (IC^M) and the objective are both unaffected by T_T^L and T_T^R.

PROOF OF COROLLARY 1:

Define the index $-j$ to be equal to L if $j = L$, and equal to R if $j = L$. Also let T_{T}^{j*} be the value of T_{T}^j such that (IC^{-j}) is just binding. I will argue that, if q and ϵ are small, the optimal punishment strategy (within the class I am considering) involves $\underline{T}_B = (\infty, \infty)$ and T_{T}^j close to T_{T}^{j*} $(j = L, R)$. The key is to keep in mind that, because of the occasional mistaken perceptions, punishment episodes are occasionally triggered along the equilibrium path (I will refer to these as "equilibrium punishments"); since q is small, the expected cost of equilibrium punishments (for the representative country) is small.

I will first argue that the optimal punishment involves $\underline{T}_B = (\infty, \infty)$. Suppose by contradiction that the optimal punishment involves $T_{\mathrm{B}}^j < \infty$ for some j. This implies that it is optimal to have a strictly positive T_T^{-j} [if it were $T_T^{-j} = 0$, the constraint (IC^j) would be binding; increasing T_T^{-j} would then imply a first-order benefit, since it would relax constraint (IC^j), and a second-order increase in the cost of equilibrium punishments]. But a punishment characterized by $T_{\mathrm{B}}^j < \infty$ and $T_T^{-j} > 0$ can be improved upon by increasing T_{B}^j and decreasing T_T^{-j}, since this can be done without affecting (IC^j) and would lower the cost of equilibrium punishments (recall that third-party sanctions by assumption involve a higher cost than bilateral ones).

Next I argue that the optimal level of T_{T}^j must be close to T_{T}^{j*} $(j = L, R)$ if q is small. Suppose first $T_{\mathrm{T}}^j > T_{\mathrm{T}}^{j*}$ for some j. Since $\underline{T}_B = (\infty, \infty)$ at an optimum, (IC^{-j}) is not binding, therefore T_{T}^j can be decreased slightly without affecting the incentive constraints, and reducing the cost of equilibrium punishments. Suppose next $T_{\mathrm{T}}^j < T_{\mathrm{T}}^{j*}$ for some j. In this case, (IC^{-j}) is clearly binding. For q small enough, it pays to increase T_{T}^j a bit, since this relaxes constraint (IC^{-j}), which implies a first-order benefit, while increasing the cost of equilibrium punishments by a second-order amount. This establishes that T_{T}^j must be close to the minimal level T_{T}^{j*} for all j.

PROOF OF PROPOSITION 4:

Consider the following strategy for government i: Charge tariffs $\underline{\tau}^M$ until a violation is observed; if a violation by country $j \neq i$ is observed (whether a substantive or a procedural violation), revert forever to the Nash tariff against that country; if a tariff violation was committed by country i itself, revert to the Nash tariffs against all the countries that were cheated against; if a procedural violation was committed by country i itself, revert to the Nash tariffs against all countries.

To check that this is an equilibrium strategy, notice that there are three possible unilateral deviations at each given time: two tariff deviations (against left and right partner), and a procedural deviation (denying access to the verifier). Let us check that a country has no incentive to commit any of these violations. The key is to establish that the only relevant incentive constraint is (IC^M) (as defined in the proof of Proposition 1), which is of course satisfied at $\underline{\tau}^M$. (1) Consider a defection on tariffs and not on procedure: since the violation will be verified, multilateral sanctions are expected for any defection. This implies that the only rel-

evant defection is against both partners. But there is no incentive to do this, because (IC^M) is satisfied at $\underline{\tau}^M$. (2) Consider a defection on tariffs *and* procedure: since the procedural violation will be observed and will be met with multilateral sanctions, a country might as well defect on both tariffs. The gain from doing this is then equal to the left-hand side of (IC^M), and the loss from cheating is equal to its right-hand side, and hence there is no incentive to violate this way. (3) Finally, it is straightforward to verify that a country has no incentive to cheat only on the procedure.

REFERENCES

Abreu, Dilip. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica*, March 1988, *56*(2), pp. 191–225.

Bagwell, Kyle and Staiger, Robert W. "A Theory of Managed Trade." *American Economic Review*, September 1990, *80*(4), pp. 779–95.

_____. "Multilateral Tariff Cooperation during the Formation of Customs Unions." *Journal of International Economics*, February 1997, *42*(1–2), pp. 91–123.

_____. "An Economic Theory of GATT." *American Economic Review,* March 1999, *89*(1), pp. 215–48.

Baldwin, Richard E. "Politically Realistic Objective Functions and Trade Policy." *Economics Letters*, August 1987, *24*(1), pp. 287–90.

Bendor, Jonathan and Mookherjee, Dilip. "Norms, Third-Party Sanctions, and Cooperation." *Journal of Law, Economics and Organization*, Spring 1990, *6*(1), pp. 33–63.

Bernheim, B. Douglas and Ray, Debraj. "Collective Dynamic Consistency in Repeated Games." *Games and Economic Behavior*, December 1989, *1*(4), pp. 295–326.

Bernheim, B. Douglas and Whinston, Michael. "Multimarket Contact and Collusive Behavior." *Rand Journal of Economics*, Spring 1990, *21*(1), pp. 1–26.

Bhagwati, Jagdish. "Departures from Multilateralism: Regionalism and Aggressive Unilateralism." *Economic Journal*, December 1990, *100*(403), pp. 1304–17.

Caplin, Andrew and Krishna, Kala. "Tariffs and the Most-Favored-Nation Clause: A Game-Theoretic Approach." *Seoul Journal of Economics*, August 1988, *1*(3), pp. 267–89.

Croome, John. *Reshaping the world trading system: A history of the Uruguay Round*. Geneva: World Trade Organization, 1995.

Dixit, Avinash. "Strategic Aspects of Trade Policy," in Truman F. Bewley, ed., *Advances in economic theory: Fifth World Congress*. Cambridge: Cambridge University Press, 1987, pp. 329–62.

Farrell, Joseph and Maskin, Eric. "Renegotiation in Repeated Games." *Games and Economic Behavior*, December 1989, *1*(4), pp. 327–60.

Finger, Michael. "GATT's Influence on Regional Arrangements," in Jaime de Melo and Arvind Panagarya, eds., *New dimensions in regional integration*. Cambridge: Cambridge University Press, 1993, pp. 128–48.

Finlayson, Jock and Zacher, Mark. "The GATT and the Regulation of Trade Barriers: Regime Dynamics and Functions." *International Organization*, Autumn 1981, *35*(4), pp. 561–602.

Goldstein, Judith. "Creating the GATT Rules: Politics, Institutions, and American Policy," in John Gerard Ruggie, ed., *Multilateralism matters*. New York: Columbia University Press, 1993, pp. 201–32.

Green, Edward and Porter, Robert. "Noncooperative Collusion Under Imperfect Information." *Econometrica*, January 1984, *52*(1), pp. 87–100.

Greif, Avner; Milgrom, Paul and Weingast, Barry. "Coordination, Commitment and Enforcement: The Case of the Merchant Guild." *Journal of Political Economy*, August 1994, *102*(4), pp. 745–76.

Hudec, Robert. *The GATT legal system and world trade diplomacy*, 2nd Ed. Salem, NH: Butterworth, 1975.

_____. *Enforcing international trade law: the evolution of the modern GATT legal system*. Salem, NH: Butterworth, 1993.

Hungerford, Thomas. "GATT: A Cooperative Equilibrium in a Noncooperative Trading Regime?" *Journal of International*

Economics, November 1991, *31*(3–4), pp. 357–69.

Jackson, John H. *World trade and the law of GATT*. Indianapolis, IN: Bobbs-Merrill, 1969.

Johnson, Harry G. "Optimum Tariffs and Retaliation." *Review of Economic Studies*, 1953–1954, *21*(2), pp. 142–53.

Kandori, Michihiro. "Social Norms and Community Enforcement." *Review of Economic Studies*, January 1992, *59*(1), pp. 63–80.

Keohane, Robert. *After hegemony*. Princeton, NJ: Princeton University Press, 1984.

Kovenoch, Dan and Thursby, Marie. "GATT, Dispute Settlement and Cooperation." *Economics and Politics*, March 1993, *4*(1), pp. 151–70.

Krasner, Stephen. *International regimes*. Ithaca, NY: Cornell University Press, 1983.

Ludema, Rodney. "Optimal International Trade Agreements and Dispute Settlement Procedures." Mimeo, University of Western Ontario, 1990.

Maggi, Giovanni. "The Role of Multilateral Institutions in International Trade Cooperation." Mimeo, Princeton University, 1996.

Martin, Lisa. "The Rational State Choice of Multilateralism," in John Gerard Ruggie, ed., *Multilateralism matters*. New York: Columbia University Press, 1993, pp. 91–121.

Milgrom, Paul; North, Douglass and Weingast, Barry. "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champaigne Fairs." *Economics and Politics*, March 1990, *2*(1), pp. 1–23.

Osborne, Martin and Rubinstein, Ariel. *Bargaining and markets*. San Diego, CA: Academic Press, 1990.

Oye, Kenneth. *Cooperation under anarchy*. Princeton, NJ: Princeton University Press, 1986.

Riezman, Raymond. "Dynamic Tariffs with Asymmetric Information." *Journal of International Economics*, May 1991, *30*(3–4), pp. 267–83.

Staiger, Robert W. "International Rules and Institutions for Trade Policy," in Gene M. Grossman and Kenneth Rogoff, eds., *Handbook of international economics*, Vol. 3. Amsterdam, North-Holland, 1995, pp. 1495–551.

Winters, L. Alan. "The Road to Uruguay." *Economic Journal*, December 1990, *100*(403), pp. 1288–303.

Yarbrough, Beth, and Yarbrough, Robert. *Cooperation and governance in international trade*. Princeton, NJ: Princeton University Press, 1992.