INTERNATIONAL ECONOMIC REVIEW
Vol. 43, No. 4, November 2002

TRADE AND DOMESTIC POLICY LINKAGE IN INTERNATIONAL AGREEMENTS*

BY JOSH EDERINGTON[1]

University of Miami

A central question in discussions of integrating negotiations over domestic policy (e.g., environmental policy or labor standards) into traditional trade agreements is the degree to which the trade policy and domestic policy provisions of an agreement should be explicitly linked. For example, should the World Trade Organization enforce domestic policy obligations with the threat of the suspension of trade concessions? This article considers the conditions under which linking trade and domestic policy agreements within a self-enforcing agreement is beneficial, and argues that the benefits of such policy linkage may be lower than is commonly thought.

1. INTRODUCTION

Under the auspices of the General Agreement on Tariffs and Trade (GATT) negotiations, the international community has made great strides in lowering tariff barriers to trade. However, as tariff barriers have fallen, attention has shifted to the use of domestic policy (e.g., environmental policy, labor standards, or competition policy) as a secondary trade barrier. Beginning with the Tokyo Round in the 1970s, regulation of domestic policies within international trade agreements has become increasingly important. Indeed, both the Ministerial Meeting in 1994 (at the close of the Uruguay Round) and the recent unsuccessful Ministerial Conference in Seattle witnessed increased demands for the World Trade Organization (WTO) to more fully address labor and environmental policies. Of primary concern in these discussions is the degree to which the trade policy and domestic policy provisions of an agreement should be explicitly linked (e.g., should the WTO enforce domestic policy obligations with the threat of the suspension of trade concessions?).

In a sense, one of the most significant episodes of policy linkage was the decision at the Uruguay Round to consolidate previous trade agreements into a common legal framework under the concept of a "single undertaking": that members adhere to both GATT obligations and to the range of previous trade pacts that had been negotiated under GATT auspices (e.g., the Tokyo Round codes and the new accords in services, investment, and intellectual property). Thus, membership in GATT is now made dependent on undertaking substantial new trade obligations,

* Manuscript received April 2000; revised November 2000, June 2001.

[1] I would like to thank Bob Staiger, Bob Baldwin, Wolfgang Keller, Jenny Minier, and two anonymous referees for helpful comments and suggestions. All remaining errors are, of course, my own.
Please address correspondence to: Josh Ederington, Department of Economics, University of Miami, P.O. Box 248126, Coral Gables, FL 33124, USA. E-mail: *ederington@miami.edu.*

and these obligations are enforced with an integrated dispute settlement procedure that permits "cross-retaliation" (i.e., failure to observe an obligation in one negotiating area could result in the suspension of concessions in another area). In this sense, the Uruguay Round seems to reflect a belief in the benefits of linking policy agreements.

Interestingly, in other areas, the WTO is more cautious about linking policy obligations. For example, GATT/WTO rules forbid the use of trade sanctions as a threat either to induce members to accept obligations or to enforce international cooperation in agreements outside of those negotiated within GATT/WTO. In addition, GATT negotiators have agreed that the use of cross-retaliation should be strictly limited.[2] Taken together, these suggest uncertainty about the degree to which obligations should be linked to the enforcement of commitments made in international negotiations.

Issue linkage was initially proposed as a means of handling asymmetries among countries (see Sebenius, 1983, Charnovitz, 1998, in the political science literature, and Folmer et al., 1993, Cesar and de Zeeuw, 1996, in the environmental economics literature). In these papers, issue linkage refers to simultaneous negotiation over multiple policy issues. Intuitively, if some countries gain on a given issue whereas others gain on a second issue, then linking the two issues may facilitate a mutually beneficial agreement.[3] An alternative definition of issue linkage concerns the case where signing an agreement covering one policy issue is made conditional on signing another agreement covering a separate policy issue (i.e., single issue agreements are ruled out). This type of linkage has been proposed as a means of handling free-rider problems in international agreements. For example, since the environmental benefits of a global environmental agreement are not excludable, some countries may choose to free-ride on such agreements. Carraro and Siniscalco (1995, 1997) show that uncommitted countries may be induced to join the environmental agreement when it is linked to a separate agreement with benefits appropriable only to signatory countries. A more recent paper by Conconi and Perroni (2002) points out that although such "tie-in restrictions" can facilitate multilateral cooperation by limiting the set of feasible objections to an agreement, they can also hinder cooperation by limiting the set of counterobjections.

In contrast to much of the previous literature, this article is concerned with the question of linkage as it regards the enforcement of the agreement. Specifically, I analyze linkage in a repeated game setting where agreements are constrained to be self-enforcing (i.e., countries balance the gain from deviating unilaterally from the agreed-upon policies against the losses from triggering a costly retaliatory episode in the future). In this article, policies are negotiated simultaneously and linkage refers specifically to the question of enforcing these negotiated policies.

[2] Article 22.3 states that, in general, retaliation should involve the suspension of concessions or obligations that affect the same sector, or other sectors under the same agreement. However, if this is not satisfactory, concessions or obligations under another agreement may be suspended (i.e., cross-retaliation).

[3] Of course monetary side-payments across countries represent a more efficient means of facilitating such cooperation in the presence of asymmetries (a point made by Abrego et al., 2001), but issue linkage will be beneficial when such side-payments are not available.

Thus, a linked agreement is defined as an agreement in which a defection in any provision of the agreement triggers a retaliatory episode over both agreements (i.e., cross-retaliation). Alternatively, a nonlinked agreement is an agreement in which retaliation is confined to the provisions where the cheating took place (i.e., cross-retaliation over policies is forbidden).[4]

The analysis of this article is most closely related to that of Spagnolo (1996), who also investigates cross-retaliation within a repeated game setting. He argues that when issues are substitutes in the government's objective function, linking issues may help to sustain policy cooperation since simultaneous punishments over multiple (substitute) issues are stronger punishments whereas simultaneous deviations are less valuable than the sum of individual deviations (see also Spagnolo 1999). However, although Spagnolo (1996) considers linkage between multiple policy *issues*, I consider linkage between multiple policy *instruments*. Specifically, I consider the case where symmetric countries are faced with a single policy issue (a terms-of-trade externality) and have access to multiple policy instruments: an efficient instrument (trade policy) and an inefficient instrument (domestic policy). In this setting, countries are not directly concerned with the domestic policies of other countries, but only indirectly to the extent that these policies have effects on competitiveness. Thus, I consider the case where a country cares about potentially lax foreign environmental standards because such standards have indirect trade consequences (e.g., placing domestic industry at a competitive disadvantage), not because of cross-border pollution concerns.

In the model, negotiations over the efficient instrument (trade policy) induce substitution toward the less-efficient instrument as a secondary means of protection and thus an efficient international agreement will cover both policy instruments (see Copeland, 1990, and Ederington, 2001). The question raised by this article is whether a linked agreement is preferable to separate nonlinked agreements. The advantage of policy linkage is that it strengthens the punishment that can be threatened to potential deviators as it is well known (see Abreu, 1988) that the optimal punishment strategy entails the strongest possible sanctions. Indeed, linkage is weakly optimal in this model. However, when punishment involves reversion to interior Nash equilibrium, this article establishes that the stronger punishment of linkage does not support a more cooperative outcome even when enforcement represents a binding constraint on the ability of countries to cooperate. Thus, linkage is not necessary to support maximal efficiency.

Intuitively, when cooperation is sustained on both policy instruments, both deviations and punishments in the linked agreement involve only the efficient instrument (trade policy). Thus, the necessary condition for cooperation on the efficient instrument is identical in the nonlinked and linked agreements. Furthermore, the additional distortions caused by deviations in the inefficient instrument (domestic policy) from the cooperative equilibrium increase the efficiency losses of such deviations. These efficiency losses decrease the gain to deviating in the inefficient instrument relative to the gain to cooperating in the inefficient instrument.

[4] The definition of linkage in this article is similar to that in Bernheim and Whinston's (1990) model of oligopolistic competition across multiple markets.

Therefore, the necessary condition for cooperation to be sustained on the inefficient instrument is less stringent than for the efficient instrument. Since neither of the conditions for cooperation to be sustained on a single instrument in the nonlinked agreement binds at the most-cooperative equilibrium, linkage is not necessary to enforce the agreement. These results suggest that trade policy sanctions are not required to support the domestic policy provisions of an international agreement (and vice versa).

However, as has been mentioned by Dixit (1987), an alternative Nash equilibrium for trade policy (and the most severe punishment threat) consists of autarky-level trade barriers. As I show in Section 5, when reversion to autarky is used as punishment within the linked agreement, the threat of (autarky) trade sanctions will be necessary to support the domestic policy provisions of the agreement since autarky cannot be supported as a subgame-perfect punishment strategy with domestic policy sanctions. Thus, in an international agreement supported by the tacit threat of reversion to autarky, policy linkage is beneficial in achieving maximal efficiency (although the severity of the autarky threat makes it somewhat less plausible than the threat of reversion to interior Nash equilibrium).

I make these points in a two-country, two-good model of trade. Section 2 lays out the basic model and solves for Nash and Pareto-efficient policies. Section 3 defines the self-enforcing agreement and the distinction between linked and nonlinked agreements. Section 4 compares the linked agreement to the nonlinked agreement with interior Nash punishment, and Section 5 repeats the analysis where autarky punishment is allowed. Finally, Section 6 considers renegotiation-proof strategies, and Section 7 concludes.

2. THE MODEL

The analysis is conducted in a two-good partial equilibrium model of trade with two symmetric countries, a home and foreign country (denoted by *). Each country chooses trade and domestic policies (import tariffs and production taxes, respectively) for its import-competing industry. In Section 5, I consider the case where countries also have access to export-sector policies.

Demand functions are symmetric with the representative linear demand function for the home country denoted by $D(p_i^d)$ where p_i^d is the local consumer price of the ith good (foreign demand $D(p_i^{d*})$ is identical and defined in terms of the foreign consumer price p_i^{d*}). Supply functions also take a linear form, with x (y) being the "natural" import good of the home (foreign) country. Specifically, in each country, the domestic supply function for the import good is given by $Q_m(p_m^s)$ where p_m^s is the local producer price of the import good. Similarly, the domestic supply function of the export good is given by $Q_e(p_e^s)$ where p_e^s is the local producer price of the export good. Comparative advantage can then be established by having $Q_e > Q_m$ for a given producer price.[5]

[5] The partial equilibrium nature of the model can be rationalized in general equilibrium terms by assuming utility functions linear in a "residual" good and separable in the other goods. Provided that the residual good is always consumed in positive amounts, the marginal utility will be fixed at one and

The home-country production tax (in the import-competing industry) is denoted by t, and τ is the specific tariff applied to imports. Likewise, t^* and τ^* denote, respectively, the production tax and import tariff choices of the foreign country. Therefore, in the home country, producer and consumer prices in the import sector (provided that trade taxes are not prohibitive) are given, respectively, by $p_x^s(p_x^w, t, \tau) \equiv p_x^w - t + \tau$ and $p_x^d(p_x^w, \tau) \equiv p_x^w + \tau$, with p_x^w denoting the "world" (untaxed) price of good x. The absence of export-sector intervention implies that prices in the home-country's export sector are given by $p_y^s = p_y^d \equiv p_y^w$. Prices in the foreign country are symmetrically defined: $p_y^{s*}(p_y^w, t^*, \tau^*) \equiv p_y^w - t^* + \tau^*$, $p_y^{d*}(p_y^w, \tau^*) \equiv p_y^w + \tau^*$, and $p_x^{s*} = p_x^{d*} \equiv p_x^w$. The important distinction between trade and domestic policies is that trade policies drive a wedge between the world price and domestic prices, whereas domestic policies drive a wedge between domestic producer and consumer prices.

Given a positive trade volume, world markets will clear (i.e., world demand will equal world supply). From this market-clearing condition, the market-clearing world price for each good can be derived: $p_x^w(\tau, t)$ and $p_y^w(\tau^*, t^*)$ and, thereby, local producer and consumer prices for the jth country: $\hat{p}_x^{sj}(t, \tau)$, $\hat{p}_x^{dj}(t, \tau)$ and $\hat{p}_y^{sj}(t^*, \tau^*)$, $\hat{p}_y^{dj}(t^*, \tau^*)$. Using these local prices, the market-clearing import (M_x) and export volume (E_y) can be calculated from excess demand. The foreign import (M_y^*) and export (E_x^*) volumes follow directly from the market-clearing condition that $M_x = E_x^*$ and $E_y = M_y^*$.

Since domestic policies (e.g., environmental policy) are commonly justified as correcting for domestic distortions (e.g., pollution), I allow for the presence of a purely domestic externality arising from the production of the import good within each country, the cost of which is given by the functions $S(Q_m) = s \cdot Q_m(p_x^s)$ and $S^*(Q_m^*) = s \cdot Q_m^*(p_y^{s*})$ where s reflects the marginal external cost of production.[6] Importantly, I assume that there are no international "spillovers" associated with this externality.

Finally, governments are assumed to maximize the sum of consumer surplus and producer surplus, net of external costs of production, and trade policy and domestic policy revenue. Welfare functions for the home country over import and export goods, respectively, are given by

(1)

$$W_x(t, \tau) \equiv \int_{\hat{p}_x^d(t,\tau)}^{1} D(t, \tau)\, d\hat{p} + \int_0^{\hat{p}_x^s(t,\tau)} Q_m(t, \tau)\, d\hat{p} + \tau \cdot M_x(t, \tau) + [t - s] \cdot Q_m(t, \tau)$$

$$W_y(t^*, \tau^*) \equiv \int_{\hat{p}_y^d(t^*,\tau^*)}^{1} D(t^*, \tau^*)\, d\hat{p} + \int_0^{\hat{p}_y^s(t^*,\tau^*)} Q_e(t^*, \tau^*)\, d\hat{p}$$

partial equilibrium analysis of the nonnumeraire sectors is appropriate. The supply functions can be derived from underlying production functions ($Q_i(L_i)$ where L_i is labor used in production of the ith good) under the assumption that labor supply is infinitely elastic at a unitary wage.

[6] The assumption that pollution arises only from production of the import-competing good (which is relaxed in Section 5) is made in order to be consistent with my assumption that countries only have access to import-sector policies.

Aggregate welfare is then given by $W(t, \tau, t^*, \tau^*) \equiv W_x(t, \tau) + W_y(t^*, \tau^*)$. Foreign welfare is defined symmetrically as the sum of welfare from import and export goods (that is, $W^*(t, \tau, t^*, \tau^*) = W_x^*(t, \tau) + W_y^*(t^*, \tau^*)$).

In the absence of an international agreement, each country sets trade taxes and production taxes to maximize national welfare, taking the policy choices of its trading partner as given. The unilaterally optimal domestic and trade policies for the home country (with optimal noncooperative foreign policies being symmetrically defined) satisfy the following first-order conditions:

$$[t - s] + \lambda_t \left[\tau - \tfrac{1}{\varepsilon(\tau,t)}\right] = 0 \tag{2}$$

$$[t - s] + \lambda_\tau \left[\tau - \tfrac{1}{\varepsilon(\tau,t)}\right] = 0 \tag{3}$$

where

$$\varepsilon(\tau, t) \equiv \frac{\partial E_x^*(t, \tau)/\partial \tau}{E_x^*(t, \tau) \cdot \left[\partial p_x^w(t, \tau)/\partial \tau\right]} > 0, \qquad \lambda_t \equiv \frac{\partial M_x(t, \tau)/\partial t}{\partial Q_m(t, \tau)/\partial t} < 0,$$

$$\lambda_\tau \equiv \frac{\partial M_x(t, \tau)/\partial \tau}{\partial Q_m(t, \tau)/\partial \tau} < 0 \qquad \text{and} \quad \lambda_t \neq \lambda_\tau$$

Viewed from a single country's perspective, there are two basic distortions in this model: a production distortion and an international trade distortion (the policy choices of each country affect the terms of trade). Although countries can use either policy instrument to restrict trade and achieve terms-of-trade benefits, solving (2) and (3) for the optimal, noncooperative trade and domestic taxes yields the standard result of welfare analysis: The first-best policy choice is to pursue terms-of-trade advantages with trade policies (set $\tau^N = 1/\varepsilon$) and counter domestic distortions with domestic policies (set $t^N = s$). Since markets for the two goods are independent and export policies are prohibited, these optimal policy choices are independent of foreign policy. Nash welfare for the home country is then given by $W(\tau = \tau^N, t = t^N, \tau^* = \tau^{N*}, t^* = t^{N*}) \equiv W^N$, with Nash welfare for the foreign country symmetrically defined and denoted by W^{N*}.

Efficient trade and domestic policies will be set to maximize joint welfare ($W + W^*$) and will serve as the natural goals toward which countries strive when they cooperate. I will focus on symmetric international agreements in which countries set common cooperative trade policies ($\tau = \tau^* = \tau^c$) and domestic policies ($t = t^* = t^c$), since common policies in a symmetric model imply that both countries share equally in the gains to cooperation. Given that countries set common cooperative policies, the symmetry of the model implies that maximization of joint welfare will be equivalent to maximization of home welfare (defined by $W(\tau = \tau^c, t = t^c, \tau^* = \tau^c, t^* = t^c) \equiv W^C(\tau^c, t^c)$). The efficient symmetric domestic ($\bar{t}^c$) and trade ($\bar{\tau}^c$) policies are a combination of free trade ($\bar{\tau}^c = 0$) and nondistortionary production taxes ($\bar{t}^c = s$). Thus, the goal of international cooperation is to achieve efficiency by (i) eliminating the terms-of-trade motivations from each country's trade policy decisions, and (ii) preventing each country from distorting its domestic policy as a secondary means of protection.[7]

3. POLICY LINKAGE

The previous section established that countries have a unilateral incentive to erect barriers to trade, but since such barriers are globally suboptimal, countries would benefit from cooperating toward free trade while retaining their nondistortionary domestic policies. However, as suggested by Copeland (1990), trade negotiations that constrain the ability of countries to pursue terms-of-trade advantages through trade policies induce substitution toward nonnegotiable secondary trade barriers such as domestic policy. Note, from (2), that as the cooperative tariff is lowered below the Nash tariff, the home country has a unilateral incentive to lower its domestic tax below the globally efficient (nondistortionary) level. Thus, any efficient agreement will entail cooperation over both trade and domestic policy.

Unfortunately, the desire to erect trade barriers does not disappear once an agreement is in place, and a critical problem faced by any international agreement is the lack of an external enforcement mechanism to ensure that the signatories to an agreement uphold their obligations. In the absence of external enforcement mechanisms, an agreement will only be viable if it is self-enforcing (i.e., member countries must view their continued cooperation to be in their own best interest). To model this enforcement issue, I employ the infinitely repeated tariff game outlined by Dixit (1987) and Bagwell and Staiger (1990), in which tacit cooperation is sustained by the credible (subgame-perfect) punishment of infinite reversion to a Nash equilibrium.[8]

3.1. *Linked Agreements.* When trade and domestic policies are linked within the agreement, then in each period, each country plays the cooperative trade and domestic policy pair (τ^c, t^c) provided that cooperative policies have been played by both countries in all previous periods. Since deviation in either trade or domestic policy will be punished by reversion to the Nash in both policies, if a country chooses to deviate from the linked agreement, it will do so in both trade and domestic policy (see Bernheim and Whinston, 1990). Note that a deviating country will select a trade and domestic tax pair (τ^D, t^D) along its best response curves, defined by Equations (2) and (3) to be $\tau^D = \tau^N = 1/\varepsilon, t^D = t^N = s$. Given the symmetry of the model, it should be apparent that all results concerning the incentive to defect can be expressed in the notation of the home country (as they will hold equivalently for the foreign country). Defining $W(\tau = \tau^D, t = t^D, \tau^* = \tau^c, t^* = t^c) \equiv W^D(\tau^c, t^c)$, the one-period gain to deviating from the agreement in both trade and domestic policy (Ω) is then given by $\Omega(\tau^c, t^c) \equiv W^D(\tau^c, t^c) - W^C(\tau^c, t^c)$.

[7] For an early analysis of the terms-of-trade justification for international cooperation, see Johnson (1953–54). In a more recent paper, Bagwell and Staiger (1999) provide a terms-of-trade interpretation of international agreements and GATT principles.

[8] Thus, I analyze the case of unrelenting trigger strategies as in Friedman (1971). Similar results can be derived if one assumes that deviation triggers reversion to the Nash equilibrium for a finite number of periods.

However, cheating on the agreement (in either trade or domestic policy) causes both countries to revert to the noncooperative policies (τ^N and t^N) in all future periods. The per-period benefit to maintaining cooperation in both policies (ω) is given by $\omega(\tau^c, t^c) \equiv W^C(\tau^c, t^c) - W^N$. Defining δ as the discount factor between periods, the discounted value of avoiding a (permanent) breakdown in cooperation is given by $[\delta/(1-\delta)]\omega(\tau^c, t^c)$. An optimal international agreement results in both countries jointly choosing trade and domestic policies to maximize the cooperative level of welfare $W^C(\tau^c, t^c)$ subject to the self-enforcement constraint

$$g(\tau^c, t^c) \equiv \Omega(\tau^c, t^c) - \frac{\delta}{(1-\delta)}\omega(\tau^c, t^c) \preceq 0 \tag{4}$$

The self-enforcement constraint (4) requires that the one-period gain to deviation must be less than the discounted value of future cooperation. Note that when countries place great value on the future (high δ), the self-enforcement constraint will not bind and the efficient symmetric trade and domestic policy pair ($\bar{\tau}^c = 0, \bar{t}^c = s$) can be supported as a self-enforcing agreement. In this article, I will focus on the situation where self-enforcement is a binding constraint on the ability of countries to cooperate. The "most-cooperative" trade and domestic policy pair ($\hat{\tau}^c, \hat{t}^c$) is then defined as the policy pair that maximizes $W^C(\tau^c, t^c)$ from among the set of enforceable policy pairs (i.e., (τ^c, t^c) such that $g(\tau^c, t^c) \leq 0$) for any δ.

3.2. *Nonlinked Agreements.* An alternative is to have two separate (nonlinked) agreements in which countries select symmetric trade and domestic taxes and credibly threaten infinite reversion to the noncooperative (Nash) equilibrium of the policy whose provision was violated. I impose a strict definition of nonlinkage in which deviations with respect to domestic (trade) policy cannot be punished by allowing countries to either rewrite the cooperative trade (domestic) policy provision or trigger punishment with respect to trade (domestic) policy.

Accordingly, the one-period gain to deviation, when a country chooses to deviate in domestic policy, will be given by $\Omega_t(\tau^c, t^c) \equiv W(\tau = \tau^c, t = t^D(\tau^c), \tau^* = \tau^c, t^* = t^c) - W^C(\tau^c, t^c)$, where $t^D(\tau^c)$ is defined by Equation (2). However, such a deviation will trigger infinite reversion to noncooperative domestic policy strategies and thus the per-period gains to cooperation in domestic policy are given by $\omega_t(\tau^c, t^c) \equiv W^C(\tau^c, t^c) - W(\tau = \tau^c, t = t^D(\tau^c), \tau^* = \tau^c, t^* = t^{D*}(\tau^c))$. Likewise, the one-period gain to deviation in trade policy will be given by $\Omega_\tau(\tau^c, t^c) \equiv W^D(\tau = \tau^D(t^c), t = t^c, \tau^* = \tau^c, t^* = t^c) - W^C(\tau^c, t^c)$, where $\tau^D(t^c)$ is given by (3), and the per-period gains to cooperation in trade policy are given by $\omega_\tau(\tau^c, t^c) \equiv W^C(\tau^c, t^c) - W(\tau = \tau^D(t^c), t = t^c, \tau^* = \tau^{D*}(t^c), t^* = t^c)$.

As in Section 3.1, an international agreement results in both countries jointly choosing trade and domestic policies to maximize $W^C(\tau^c, t^c)$ subject to the self-enforcement constraints. For the agreement to be viable, neither country can have an incentive to deviate in any policy, which entails that (4) is satisfied as well as the constraints that

$$\Omega_\tau(\tau^c, t^c) - \tfrac{\delta}{1-\delta}\omega_\tau(\tau^c, t^c) \leq 0 \tag{5}$$

$$\Omega_t(\tau^c, t^c) - \tfrac{\delta}{1-\delta}\omega_t(\tau^c, t^c) \leq 0 \tag{6}$$

In the linked agreement, deviation in a single policy was never profitable: Since any deviation triggered the linked punishment, if a country were to deviate, it would deviate in both trade and domestic policy. Therefore, the analogues to conditions (5) and (6) were suppressed in Section 3.1, since they do not bind when (4) is satisfied. However, in a nonlinked agreement, it is no longer necessarily the case that (5) and (6) are slack constraints. Therefore, it is possible that a linked agreement can support a more cooperative equilibrium than its nonlinked counterpart. Accordingly, assessing the benefits of policy linkage reduces to assessing whether an agreement that does not permit cross-retaliation can support the most-cooperative outcome supportable by a linked agreement.

4. LINKAGE VERSUS NONLINKAGE IN SELF-ENFORCING AGREEMENTS

To compare the linked agreement to the nonlinked agreement, I will first solve for the most-cooperative equilibrium within the linked agreement and then check whether the same cooperative equilibrium can be supported by separate nonlinked agreements. Cooperation over trade and domestic policies in a linked agreement is discussed in more detail in Ederington (2001), so the discussion of the most-cooperative equilibrium that follows will be fairly brief. First, I establish the following lemma.

LEMMA 1. *In the self-enforcing agreement, countries will cooperate to set nondistortionary domestic policies* ($\hat{t}^c = s$).

PROOF. See Appendix A.1 and Ederington (2001).

The intuition behind this result rests on first-best principles. The underlying reason that countries want to defect from the international agreement is trade-related (note that the international externality in the model is a terms-of-trade externality) and allowing countries protection in trade policy will be the most efficient means of countering this incentive to deviate. Therefore, a cooperative agreement will set domestic policy at its efficient level while relaxing cooperation in trade policy. Given $\hat{t}^c = s$, the gain to cooperation, $\omega(\tau^c, \hat{t}^c = s)$, is monotonically decreasing in τ^c for $\tau^c < \tau^N$. Thus, within the linked agreement, countries will coordinate over the enforceable tariffs to choose the "most-cooperative" trade tax: the smallest nonnegative tariff that satisfies the self-enforcement constraint ($\hat{\tau}^c(\delta)$).

As the most-cooperative policy pair supported in a linked agreement is given by $\hat{t}^c = s$ and $\hat{\tau}^c(\delta)$, can the same most-cooperative policy pair be supported in separate, nonlinked agreements? First, consider the question of whether self-enforcement with respect to a deviation in trade policy alone (condition (5)) is satisfied, at $\hat{\tau}^c(\delta)$ and $\hat{t}^c = s$. Given $\hat{t}^c = t^N = s$, any deviation from the linked agreement involves only trade policy (as the most efficient instrument), and thus the short-run gains to deviating from the linked agreement are equal to the short-run gains to deviating from the nonlinked trade agreement (i.e., $\Omega(\hat{\tau}^c, \hat{t}^c) = \Omega_\tau(\hat{\tau}^c, \hat{t}^c)$). Also, note that within both the linked and nonlinked agreements, since $\hat{t}^c = t^N = s$, any deviation from cooperative trade policies is punished by infinite reversion to the Nash tariff equilibrium (and thus, $\omega(\hat{\tau}^c, \hat{t}^c) = \omega_\tau(\hat{\tau}^c, \hat{t}^c)$). Therefore, given that $\hat{\tau}^c(\delta)$ and $\hat{t}^c = s$ satisfy (4), it is straightforward to derive that (5) will also be satisfied at $\hat{\tau}^c(\delta)$ and $\hat{t}^c = s$.[9]

Next, consider whether self-enforcement with respect to deviation in domestic policy alone (condition (6)) is satisfied, at $\hat{\tau}^c(\delta)$ and $\hat{t}^c = s$. Although the linked agreement punishes any deviation in domestic policy with trade policy sanctions (infinite reversion to the Nash tariff equilibrium), the nonlinked agreement can only punish deviations in domestic policy with domestic policy sanctions (infinite reversion to Nash domestic policies). Thus, the final question is whether countries would choose to deviate from the domestic policy provisions of the agreement if they were not linked to the trade policy provisions of the agreement. This question is of practical importance given recent calls for the international community to support international agreements over domestic policy (particularly environmental policy and labor standards) with the threat of trade sanctions. The implicit assumption in these demands is that the threat of reversion to the noncooperative equilibrium in domestic policy alone will not carry enough weight to support an international agreement over domestic policy.

A sufficient condition for (6) to be satisfied when (4) holds (i.e., for the nonlinked agreement to be able to support any cooperative policy pair $(\hat{\tau}^c(\delta), \hat{t}^c = s)$ supported by a linked agreement) is that[10]

$$\frac{\omega_t(\hat{\tau}^c, \hat{t}^c = s)}{\omega_\tau(\hat{\tau}^c, \hat{t}^c = s)} \geq \frac{\Omega_t(\hat{\tau}^c, \hat{t}^c = s)}{\Omega_\tau(\hat{\tau}^c, \hat{t}^c = s)} \tag{7}$$

Condition (7) states that, for a nonlinked agreement to be equivalent to a linked agreement, the relative loss to being punished in domestic policy (as opposed to being punished in trade policy) must outweigh the relative gain to deviating in domestic policy (as opposed to deviating in trade policy). If this condition is not satisfied, then the nonlinked agreement is not self-enforcing, and linking the two policies would be preferred.

First, note that the net gain to deviation in either policy is equal to the terms-of-trade gains to restricting trade minus the efficiency losses of deviating from

[9] The equivalence between a joint deviation in both policies and deviation in trade policy alone does depend upon the assumption of a constant marginal external cost to production. With nonlinear externalities, deviations in trade policy from the agreement may affect this external cost and thus create production distortions (since the cooperative production tax does not change). However, (5) will still be satisfied as these production distortions will decrease the incentive to deviate in trade policy alone.

[10] This condition is derived by noting that (5) holds with equality at the most-cooperative equilibrium. Solving this equality for $\delta/(1-\delta)$ and substituting in (6) yields (7).

cooperative policies (i.e., $\Omega_\tau = G_\tau - \xi_\tau^D$ and $\Omega_t = G_t - \xi_t^D$, where G_τ (G_t) represent the terms-of-trade gains to deviation in trade (domestic) policy and ξ_τ^D (ξ_t^D) represents the efficiency losses to deviation in trade (domestic) policy). Since trade policy is the most efficient means of restricting trade (in pursuit of terms-of-trade gains), the gain to deviating in trade policy will outweigh the gain to deviating in domestic policy (i.e., $\Omega_\tau(\hat{\tau}^c, \hat{t}^c) \geq \Omega_t(\hat{\tau}^c, \hat{t}^c)$).

Next, note that, given the symmetry of the model, the terms-of-trade gains to reverting to Nash policies cancel out and countries are left with only the efficiency losses caused by deviating from cooperative policies. Thus, the per-period gain to cooperation is equivalent to the value of avoiding those efficiency losses (i.e., $\omega_\tau = \xi_\tau^N$ and $\omega_t = \xi_t^N$ where ξ_τ^N (ξ_t^N) represents the efficiency losses of reversion to noncooperative trade (domestic) policies). Since noncooperative trade policies (τ^N, τ^{N*}) represent greater deviations from the cooperative equilibrium than noncooperative domestic policies (t^D, t^{D*}), the loss to reverting to noncooperative trade policies is greater than the loss to reverting to noncooperative domestic policies (i.e., $\omega_\tau(\hat{\tau}^c, \hat{t}^c) \geq \omega_t(\hat{\tau}^c, \hat{t}^c)$.

It should be apparent from this discussion that, while linking the two agreements leads to greater punishment when a member country deviates in domestic policy, the domestic policy provisions will require less support than the trade policy provisions of the agreement (and thus, may not require policy linkage to be enforceable). To find the overall effect of policy linkage, note that, since best-response policies are noncooperative policies, $\xi_\tau^D = \xi_\tau^N$ and $\xi_t^D = \xi_t^N$. Thus, substituting for Ω and ω, one can derive that (7) is satisfied provided that $G_\tau/\xi_\tau \geq G_t/\xi_t$. Given the inefficiency of domestic policy as a means of pursuing terms-of-trade gains, this inequality is satisfied and the most-cooperative equilibrium ($\hat{\tau}^c(\delta), \hat{t}^c = s$) can be supported with nonlinked agreements.

Intuitively, the additional production distortions caused by deviations in domestic policy from the cooperative equilibrium decrease the gain to deviating in domestic policy and increase the gain to cooperation in domestic policy. Consequently, when self-enforcement with respect to trade policy is satisfied, then self-enforcement with respect to a deviation in domestic policy will also be satisfied (and the linked and nonlinked agreements will be equivalent).

This intuition is verified in Proposition 1.

PROPOSITION 1. *In a self-enforcing agreement that is enforced by the implicit threat of reversion to interior Nash equilibria, policy linkage will not be beneficial (i.e., the most-cooperative equilibrium, $\hat{t}^c = s$, and $\tau^c = \hat{\tau}^c(\delta)$, can be supported in both the linked and nonlinked agreements).*

PROOF. See Appendix A.2.

Note, finally, that even in the nonlinked agreement, the optimal cooperative domestic tax, t^c, is not a direct function of the discount factor δ. This may seem somewhat surprising; however, recall that the incentive to deviate from $\hat{t}^c = s$ is a function of $\hat{\tau}^c(\delta)$, which is in turn a function of δ. Intuitively, a welfare-maximizing agreement, by relaxing cooperation to ensure that neither country has an incentive to deviate with trade policy, is also implicitly assuring that neither country will have an incentive to deviate with domestic policy.

5. EXPORT POLICIES AND AUTARKY PUNISHMENT

In the previous section, it was established that separate (nonlinked) agreements can support the same most-cooperative equilibrium as a linked agreement. However, the preceding analysis assumes that punishment consists of reversion to *interior* Nash equilibria. Although Nash reversion is the most common subgame-perfect punishment strategy in the literature on self-enforcing agreements, it is not always an optimal strategy in the sense of being the most severe punishment (see Abreu, 1988). Indeed, as mentioned by Dixit (1987) and Ludema (2001), the most severe punishment threat within an international trade agreement is infinite reversion to autarky. This is of potential importance since autarky-level trade barriers can be supported as a Nash equilibrium in the case where countries have access to export-sector trade policies, whereas no equivalent (autarky-level) Nash equilibrium exists for domestic policy retaliation.[11] Thus, the most severe punishment threat (infinite reversion to autarky) can only be supported as a subgame-perfect punishment in the case of trade policy retaliation. As I show in this section, when autarky reversion is used as punishment within the linked agreement, the threat of (autarky) trade sanctions may be necessary to support the domestic policy provisions of the agreement.

This result can be illustrated by expanding the model of Section 2 to include export-sector policies. To simplify the analysis, I assume explicit functional forms for the demand and supply functions. Thus, in the home country, demand for good i is given by $D(p_i^d) = 1 - p_i^d$, and it is defined identically in the foreign country ($D^*(p_i^{d*}) = 1 - p_i^{d*}$). Supply functions are chosen to ensure that x (y) is the "natural" import good of the home (foreign) country. Specifically, in each country, the domestic supply function for the import good is given by $Q_m(p_m^s) = p_m^s$ whereas the domestic supply function of the export good is given by $Q_e(p_e^s) = \alpha + p_e^s$, where $\alpha \in [0, 2/3]$.

The home country now chooses production taxes on both the import-competing (t_m) and export sector (t_e). In addition, the home country chooses trade taxes on both sectors where τ_m is the (nonnegative) specific tariff applied to imports and τ_e is the (nonnegative) specific export tax. Likewise, t_m^*, t_e^*, τ_m^*, and τ_e^* denote production taxes and trade taxes for the foreign country. Therefore, in the home country, producer and consumer prices in the import sector (provided that trade taxes are not prohibitive) are given, respectively, by $p_x^s(p_x^w, t_m, \tau_m) \equiv p_x^w - t_m + \tau_m$ and $p_x^d(p_x^w, \tau_m) \equiv p_x^w + \tau_m$, whereas prices in the home country's export sector are given by $p_y^s(p_y^w, t_e, \tau_e) \equiv p_y^w - t_e - \tau_e$ and $p_y^d(p_y^w, \tau_e) \equiv p_y^w - \tau_e$. Prices in the foreign country are symmetrically defined.

[11] Although countries can set their domestic policies jointly to achieve autarky, doing so is not a Nash equilibrium and thus does not represent a credible, subgame-perfect threat (note that if one country is setting domestic taxes so as to achieve autarky, the other country's best response is to play the nondistortionary tax).

The global market-clearing condition yields world prices for each good: $p_x^w(\tau_m, t_m, \tau_e^*, t_e^*)$ and $p_y^w(\tau_e, t_e, \tau_m^*, t_m^*)$.[12] Using these world prices, local producer and consumer prices for each country can be expressed as functions of trade and domestic policy, and the market-clearing import and export volume can be calculated from excess demand. Finally, I assume the existence of a (symmetric) external cost to production in each sector of the economy, where $s \in [0, 2/3]$ reflects the marginal external cost of production. Thus, welfare for the home country on its import and export goods respectively is given by

$$\text{(8)} \quad \begin{aligned} W_x\big(t_m, \tau_m, \tau_e^* t_e^*\big) &\equiv \int_{\hat{p}_x^d}^1 D\big(\hat{p}_x^d\big)\, d\hat{p} + \int_0^{\hat{p}_x^s} Q_m\big(\hat{p}_x^s\big)\, d\hat{p} + \tau_m \cdot M_x\big(\hat{p}_x^d, \hat{p}_x^s\big) \\ &\quad + [t_m - s] \cdot Q_m\big(\hat{p}_x^s\big) \\ W_y\big(t_e, \tau_e, \tau_m^* t_m^*\big) &\equiv \int_{\hat{p}_y^d}^1 D\big(\hat{p}_y^d\big)\, d\hat{p} + \int_0^{\hat{p}_y^s} Q_e\big(\hat{p}_y^s\big)\, d\hat{p} + \tau_e \cdot E_y\big(\hat{p}_y^d, \hat{p}_y^s\big) \\ &\quad + [t_e - s] \cdot Q_e\big(\hat{p}_y^s\big) \end{aligned}$$

Foreign welfare is defined symmetrically as the sum of welfare from its import and export goods (i.e., $W^* = W_x^*(t_m, \tau_m, t_e^*, \tau_e^*) + W_y^*(t_e, \tau_e, t_m^*, \tau_m^*)$).

In the absence of an international agreement, each country sets trade taxes and production taxes to maximize national welfare, taking the policy choices of its trading partner as given. Taking derivatives of the welfare functions yields unilaterally optimal trade and domestic policies for the home country (τ_m^D, t_m^D, t_e^D, and τ_e^D) given by

$$\text{(9)} \quad \tau_m^D = \tfrac{1}{6}\big(\alpha - 2s + 3t_m - t_e^* - 2\tau_e^*\big), \quad \tau_e^D = \tfrac{1}{6}\big(\alpha + 2s - 3t_e + t_m^* - 2\tau_m^*\big)$$

$$\text{(10)} \quad t_m^D = \tfrac{1}{7}\big(6s - \alpha + 6\tau_m + t_e^* + 2\tau_e^*\big), \quad t_e^D = \tfrac{1}{7}\big(6s + \alpha - 6\tau_e + t_m^* - 2\tau_m^*\big)$$

Best-response functions for the foreign country are defined symmetrically and the intersection of the reaction functions defines the interior Nash equilibrium:

$$\text{(11)} \qquad \tau_m^N = \tau_e^N = \tau_m^{N*} = \tau_e^{N*} = \tfrac{\alpha}{8}, \quad t_m^N = t_e^N = t_m^{N*} = t_e^{N*} = s$$

One possibility is for cooperation within the linked agreement to be sustained by infinite reversion to the above interior Nash equilibrium. If this is the case, then it is direct to verify that the results of Section 5 hold, and the linked agreement is functionally equivalent to the nonlinked agreement (i.e., linkage does not aid in supporting a more-cooperative outcome).

However, consider the case where the foreign country sets nondistortionary domestic taxes ($t_m^* = t_e^* = s$) and prohibitive trade taxes ($\tau_m^* = \tau_e^* \geq \alpha/2$). Given

[12] Solving for these world prices yields $p_x^w = [2 - \alpha - 2\tau_m + 2\tau_e^* + t_m + t_e^*]/4$ and $p_y^w = [2 - \alpha - 2\tau_m^* + 2\tau_e + t_m^* + t_e]/4$.

these prohibitive foreign trade barriers, any home trade barrier will generate autarky, and thus home country welfare is independent of its trade taxes (thus any $\tau_m \in [0, \infty]$ and $\tau_e \in [0, \infty]$ will represent best responses to foreign autarky taxes). Similar calculations apply for the foreign country and thus one can derive a second Nash equilibrium corresponding to autarky:[13]

$$\tau_m^A = \tau_e^A = \tau_m^{A*} = \tau_e^{A*} \geq \tfrac{\alpha}{2}, \quad t_m^A = t_e^A = t_m^{A*} = t_e^{A*} = s \tag{12}$$

Hence, there is a second possible subgame-perfect punishment scheme for the linked agreement involving infinite reversion to this autarkic equilibrium. In what follows, I focus on a symmetric, self-enforcing agreement in which countries set common cooperative trade policies ($\tau_m = \tau_e = \tau_m^* = \tau_e^* = \tau^c$) and domestic policies ($t_m = t_e = t_m^* = t_e^* = t^c$) to maximize the cooperative level of welfare, $W^C(\tau^c, t^c)$, subject to the self-enforcement constraint.[14] If a country were to deviate from these common policies, it would do so by selecting unilaterally optimal trade and/or domestic policies defined by Equations (9) and (10). The one-period gain to deviating in both policies is given by $\Omega(\tau^c, t^c) \equiv [W_x(\tau_m^D, t_m^D, \tau^c, t^c) + W_y(\tau_e^D, t_e^D, \tau^c, t^c)] - W^C(\tau^c, t^c)$. Assume that such a defection from the agreement (in both policies) triggers infinite reversion to the autarkic Nash equilibrium defined by (12). Let $\omega(\tau^c, t^c) = W^A - W^C(\tau^c, t^c)$ denote the per-period benefit to maintaining cooperation, where W^A denotes the autarky level of welfare. Maximizing $W^C(\tau^c, t^c)$ subject to the self-enforcement constraint (4) yields most-cooperative policies given by $\hat{\tau}^c(\delta)$ and $\hat{t}^c = s$, where $\hat{\tau}^c(\delta)$ is the lowest common tariff level that satisfies the self-enforcement constraint. Letting $\hat{t}^c = s$, it can be shown, from (8) and the definitions of $\Omega(\tau^c, t^c)$ and $\omega(\tau^c, t^c)$, that

$$\Omega(\hat{\tau}^c, \hat{t}^c) = \tfrac{1}{24}[\alpha^2 + 64(\hat{\tau}^c)^2 - 16\alpha\hat{\tau}^c], \quad \omega(\hat{\tau}^c, \hat{t}^c) = \tfrac{1}{32}[8\alpha^2 - 64(\hat{\tau}^c)^2] \tag{13}$$

Again, the question is whether these most-cooperative equilibria ($\hat{\tau}^c(\delta), \hat{t}^c = s$) can also be supported by separate (nonlinked) agreements that forbid cross-retaliation. As in the previous section, such an equilibrium can be supported in the nonlinked agreement if condition (7) is satisfied. When the home country deviates in domestic policy alone, the one-period gain to that deviation is given by $\Omega_t(\tau^c, t^c) \equiv [W_x(\tau^c, t_m^D, \tau^c, t^c) + W_y(\tau^c, t_e^D, \tau^c, t^c)] - W^C(\tau^c, t^c)$, where t_m^D and t_e^D represent unilaterally optimal domestic policies for the home county (defined by (10)). In the nonlinked agreement, such a deviation triggers infinite reversion to the noncooperative equilibrium in domestic policy and the per-period benefit to maintaining cooperation in domestic policy is given by $\omega_t(\tau^c, t^c) \equiv$

[13] As noted by Dixit (1987), if one interprets trade policy as simply a wedge between the world price and the local price (thus τ_m would be a tax if the good is imported and a subsidy if the good is exported), then autarky would not be a possible equilibrium.

[14] The analysis would not change if one assumed that cooperative policies were not similar across sectors. However, it seems natural to focus on a common protection level given the symmetry across sectors, and doing so greatly simplifies the analysis.

$W^C(\tau^c, t^c) - [W_x(\tau^c, t_m^D, \tau^c, t_e^{D*}) + W_y(\tau^c, t_e^D, \tau^c, t_m^{D*})]$.[15] Letting $\hat{t}^c = s$, I can derive, from (8) and the definitions of $\Omega_t(\tau^c, t^c)$ and $\omega_t(\tau^c, t^c)$, that

$$\text{(14)}\quad \Omega_t(\hat{\tau}^c, \hat{t}^c) = \tfrac{1}{56}[\alpha^2 + 64(\hat{\tau}^c)^2 - 16\alpha\hat{\tau}^c], \quad \omega_t(\hat{\tau}^c, \hat{t}^c) = \tfrac{1}{64}[\alpha^2 - 64(\hat{\tau}^c)^2]$$

Comparing (13) and (14), one can verify that (7) is no longer satisfied and, thus, the nonlinked agreement cannot support the most-cooperative equilibrium of such a linked agreement (i.e., if cross-retaliation were forbidden, then each country would have an incentive to deviate from the agreement in domestic policy). Therefore, with autarky trade sanctions, policy linkage is beneficial in achieving the most-cooperative equilibrium. This result is summarized in Proposition 2.

PROPOSITION 2. *In an international agreement enforced by the threat of autarky trade sanctions, policy linkage may be beneficial (as otherwise, countries may have an incentive to defect from the domestic policy provisions of the agreement).*

Proposition 2 is of interest because it suggests that the gain from the use of trade sanctions as a means of supporting the domestic policy provisions of an agreement lies not in the efficiency of trade sanctions per se, but rather in the ability of trade sanctions to support (as a subgame-perfect punishment scheme) infinite reversion to autarky, given that there is no equivalent autarky equilibrium for retaliatory domestic policies.

However, the use of infinite reversion to autarky as a credible punishment threat is problematic. As mentioned by both Dixit (1987) and Ludema (2001), because the equilibrium strategies that bring about autarky are weakly dominated, the autarky equilibrium is unstable (i.e., chance deviations in economic conditions or an inability to completely eliminate trade will cause the autarky equilibrium to cease to exist). For this reason, many authors assume away the autarkic equilibrium. In addition, the very severity of autarky makes it less plausible as a punishment threat. For example, one can assume (as in McCutcheon, 1997) that countries can renegotiate the original agreement upon triggering the punishment phase, but that such renegotiation is costly. In this case, renegotiation will occur if the gain to renegotiation (i.e., avoiding the punishment phase) outweighs the costs of such renegotiation, implying an upper bound to the strength of punishment (i.e., $\delta \cdot \omega(\tau^c, t^c)/(1-\delta) \leq \bar{\omega}$). Provided the costs of renegotiation are sufficiently small, this renegotiation constraint will bind (i.e., $\delta \cdot \omega/(1-\delta) = \delta \cdot \omega_t/(1-\delta) = \bar{\omega}$). Given the inefficiency of domestic policy as a means of deviating from the agreement (i.e., $\Omega_t \leq \Omega$) it should be apparent that (7) will be satisfied, implying that the nonlinked agreement will again be equivalent

[15] In this section, I am implicitly comparing the relative punishment threat of infinite autarky reversion to infinite reversion to Nash domestic policies. One could consider alternative subgame-perfect punishment schemes (as in Abreu, 1988) for domestic policy retaliation; however, the severity of domestic policy retaliation does not affect the basic point that the ability of trade policy to support the minmax payoff as a subgame-perfect equilibrium provides an advantage to trade policy retaliation in sustaining maximal collusion.

to the linked agreement. In the following section, I verify that the results of this article also apply to more complicated renegotiation-proof strategies.

6. RENEGOTIATION-PROOF STRATEGIES

In the previous sections, it was established that (unless autarkic reversion can be maintained as a punishment strategy) policy linkage was not necessary to maintain maximal collusion. However, the punishment strategies used are somewhat problematic as they are sensitive to concerns about renegotiation. Specifically, it is well known that on triggering the punishment phase, countries will have an incentive to renegotiate to achieve the original agreement, and this renegotiation will reduce the amount of punishment with which potential deviators can be threatened. Thus, in this section, I consider renegotiation-proof strategies along the lines of van Damme (1989). Within the context of the partial equilibrium model where home and foreign policies are independent, I show that the results of the previous sections remain valid under such renegotiation-proof strategies.

Consider the following strategy profile of van Damme (1989): Each country plays cooperatively as long as the other country plays cooperatively. If country i defects in period t (and country k does not) then country k will defect until country i again plays cooperatively. As soon as country i has repented by playing cooperatively, country k forgives the initial defection and returns to playing cooperatively.[16]

First, let us evaluate the conditions under which the linked agreement is self-enforcing. As before, define W^C as the cooperative level of welfare, W^D as the level of welfare achieved by deviating from the agreement, and W^N as the Nash level of welfare. Finally, define $W^{\sim D}$ as the welfare for the home country when it is repenting by playing cooperative policies whereas the foreign country punishes by playing unilaterally optimal policies (i.e., $W^{\sim D} = W(\tau = \tau^c, t = t^c, \tau^* = \tau^N, t^* = t^N)$). For the home country, the conditions for the above strategy profile to be self-enforcing are given by[17]

$$\frac{1}{1-\delta} W^C \geq W^D + \delta\left[W^{\sim D} + \frac{\delta}{1-\delta} W^C\right] \tag{15}$$

$$W^{\sim D} + \frac{\delta}{1-\delta} W^C \geq W^N + \delta\left[W^{\sim D} + \frac{\delta}{1-\delta} W^C\right] \tag{16}$$

The first condition states that neither country has an incentive to deviate from cooperative policies, whereas the second condition states that the deviating country will be willing to "repent" by playing cooperative policies after triggering the punishment phase.

Next, one can derive conditions for the environmental agreement to be self-enforcing if it is not linked to the free-trade agreement. As before, define W_t^D as

[16] Note that autarkic reversion does not play a role in van Damme punishment strategies, since the autarky payoff is dominated by other payoffs and thus would be renegotiated.

[17] There is a third condition that the nondeviating country does not wish to deviate from the punishment strategy, but this condition is satisfied trivially. Likewise, the self-enforcement conditions for the foreign country hold symmetrically.

the level of welfare achieved by the home country by deviating from only the cooperative environmental policy, and W_t^N as the level of welfare when both countries are playing unilaterally optimal environmental policies. When the environmental agreement is not linked to the trade agreement, then the deviating country can repent for deviating in environmental policy by playing the cooperative environmental policy whereas the foreign country plays the unilaterally optimal environmental policy. Thus, define $W_t^{\sim D}$ as welfare for the home country when it is repenting for a past deviation (i.e., $W_t^{\sim D} = W(\tau = \tau^c, t = t^c, \tau^* = \tau^c, t^* = t^D(\tau^c))$. For the home country, the conditions for the renegotiation-proof strategy profile to be self-enforcing are given by

$$\frac{1}{1-\delta} W^C \geq W_t^D + \delta\left[W_t^{\sim D} + \frac{\delta}{1-\delta} W^C\right] \tag{17}$$

$$W_t^{\sim D} + \frac{\delta}{1-\delta} W^C \geq W_t^N + \delta\left[W_t^{\sim D} + \frac{\delta}{1-\delta} W^C\right] \tag{18}$$

Compare the conditions for cooperative strategies to be self-enforcing in the linked and nonlinked agreements (i.e., conditions (15) and (17)) at the global optimum. A sufficient condition for (17) to be satisfied when (15) holds (i.e., for the nonlinked agreement to be equivalent to the linked agreement) is for[18]

$$\frac{W_t^D - W^C}{W^C - W_t^{\sim D}} \geq \frac{W^D - W^C}{W^C - W^{\sim D}} \tag{19}$$

Next, compare the conditions for repentance to be self-enforcing in the linked and nonlinked agreements (i.e., conditions (16) and (18)) at the global optimum). A sufficient condition for (18) to be satisfied when (16) holds (i.e., for the nonlinked agreement to be equivalent to the linked agreement) is for

$$\frac{W_t^N - W_t^{\sim D}}{W^C - W_t^{\sim D}} \geq \frac{W^N - W^{\sim D}}{W^C - W^{\sim D}} \tag{20}$$

Thus, the nonlinked agreement will be able to support any cooperative policy pair $(\hat{\tau}^c(\delta), \hat{t}^c = s)$ supported by a linked agreement if both (19) and (20) are satisfied. However, one can take advantage of the independence of home and foreign policies (i.e., $W = W_x(\tau, t) + W_y(\tau^*, t^*)$), to rewrite *both* (19) and (20) as being equivalent to condition (7). Since we have already established that (7) holds, I can state Proposition 3.

PROPOSITION 3. *In a self-enforcing agreement that is enforced by renegotiation-proof punishment strategies along the lines of van Damme* (*1989*), *policy linkage will not be beneficial* (*i.e., the most-cooperative equilibrium, $\hat{t}^c = s$ and $\tau^c = \hat{\tau}^c(\delta)$, can be supported in both the linked and nonlinked agreements*).

[18] This condition is derived by noting that (15) holds with equality at the most-cooperative equilibrium. Solving this equality for δ and substituting in (17) yields (19).

7. CONCLUSION

This article has investigated the question of whether policy linkage is justified within a setting of perfect information and a single international externality (countries are large and their policy choices affect the world price of the good). In this setting, the purpose of cooperation is to eliminate the "beggar-thy-neighbor" terms-of-trade incentives from the policy choices of each country. Because trade policy is a more efficient means of pursuing these terms-of-trade gains, countries value cooperation over trade policy more highly (i.e., they fear being punished by reversion to the Nash equilibrium in trade policy more than reversion to the Nash equilibrium in domestic policy). This appears to suggest that policy linkage could support a more cooperative outcome as countries could punish deviations in domestic policy with the stronger punishment of reversion to the Nash equilibrium in trade policy. This article argues that the benefits of "policy linkage" may be lower than is commonly thought, since policy linkage is only beneficial in those agreements that are supported by the implicit threat of autarky trade sanctions.

APPENDIX

A.1. *Proof of Lemma* 1. The most-cooperative trade and domestic policy pair $(\hat{\tau}^c, \hat{t}^c)$ is defined as the policy pair that maximizes the per-period gain to cooperation from among the set of enforceable policy pairs (i.e., (τ^c, t^c) such that $g(\tau^c, t^c) \leq 0$) for any δ. From the first-order conditions of this constrained maximization, one derives the following necessary condition that must be satisfied by $(\hat{\tau}^c, \hat{t}^c)$

$$\text{(A.1)} \qquad \frac{\partial \omega / \partial t^c}{\partial \omega / \partial \tau^c} = \frac{\partial \Omega / \partial t^c}{\partial \Omega / \partial \tau^c}$$

Taking derivatives of ω with respect to τ^c and t^c, I derive that

$$\text{(A.2)} \qquad \frac{\partial \omega / \partial t^c}{\partial \omega / \partial \tau^c} = -\frac{\tau^c(\partial M_x / \partial t^c) - (t^c - s)(\partial Q_m / \partial t^c)}{\tau^c(\partial M_x / \partial \tau^c) - (t^c - s)(\partial Q_m / \partial \tau^c)}$$

Taking derivatives of Ω with respect to τ^c and t^c, I derive that

$$\text{(A.3)} \qquad \frac{\partial \Omega / \partial t^c}{\partial \Omega / \partial \tau^c} = -\frac{[\tau^c - 1/\varepsilon(\tau^c, t^c)]\partial M_x / \partial t^c - (t^c - s)(\partial Q_m / \partial t^c)}{[\tau^c - 1/\varepsilon(\tau^c, t^c)]\partial M_x / \partial \tau^c - (t^c - s)(\partial Q_m / \partial \tau^c)}$$

Substituting (A.2) and (A.3) into (A.1) yields the unique locus of tangency points along which the most-cooperative policy pair $(\hat{\tau}^c, \hat{t}^c)$ must lie. This locus is given by $t^c = s$, since

$$\text{(A.4)} \qquad \left.\frac{\partial \omega / \partial \tau^c}{\partial \omega / \partial t^c}\right|_{t^c = s} = \frac{\partial M_x / \partial \tau^c}{\partial M_x / \partial t^c} = \left.\frac{\partial \Omega / \partial \tau^c}{\partial \Omega / \partial t^c}\right|_{t^c = s}$$

This completes the proof. ∎

A.2. *Proof of Proposition* 1. Using Taylor series expansion, one can express the relative gain to deviating in trade policy versus the gain to deviating in domestic policy as a function of first- and second-order derivatives

$$\text{(A.5)}\qquad \frac{\Omega_t}{\Omega_\tau} = \frac{\frac{\partial W_x^C}{\partial t^c}[t^D - t^c] + \frac{1}{2}\frac{\partial^2 W_x^C}{\partial (t^c)^2}[t^D - t^c]^2}{\frac{\partial W_x^C}{\partial \tau^c}[\tau^N - \tau^c] + \frac{1}{2}\frac{\partial^2 W_x^C}{(\partial \tau^c)^2}[\tau^N - \tau^c]^2}$$

Likewise, one can express the loss to being punished by a reversion to the Nash equilibrium in trade policy versus the loss to reversion to the Nash in domestic policy as

$$\text{(A.6)}\qquad \frac{-\omega_t}{-\omega_\tau} = \frac{\frac{\partial\left[W_x^C + W_y^C\right]}{\partial t^c}[t^D - t^c] + \frac{1}{2}\frac{\partial^2\left[W_x^C + W_y^C\right]}{\partial (t^c)^2}[t^D - t^c]^2}{\frac{\partial\left[W_x^C + W_y^C\right]}{\partial \tau^c}[\tau^N - \tau^c] + \frac{1}{2}\frac{\partial^2\left[W_x^C + W_y^C\right]}{\partial (\tau^c)^2}[\tau^N - \tau^c]^2}$$

Note that, since $\hat{\tau}^c = s$, first derivatives will only capture the effect of each policy on the trade distortion (i.e., on trade volume or the world price)

$$\text{(A.7)}\qquad \frac{\partial W_x^c/\partial \tau^c}{\partial W_x^C/\partial t^c} = \frac{\partial\left[W_x^C + W_y^C\right]/\partial \tau^c}{\partial\left[W_x^C + W_y^C\right]/\partial t^c} = \frac{\partial p^w/\partial \tau^c}{\partial p^w/\partial t^c}$$

Taking second derivatives of the respective welfare functions

$$\text{(A.8)}\qquad \frac{\partial^2 W_x^C}{\partial (\tau^c)^2} = \frac{\partial M_x}{\partial \tau^c} - \frac{\partial M_x}{\partial \tau^c}\frac{\partial p^w}{\partial \tau^c}, \qquad \frac{\partial^2 W_x^C}{\partial (t^c)^2} = \frac{\partial Q_m}{\partial t^c} - \frac{\partial M_x}{\partial t^c}\frac{\partial p^w}{\partial t^c}$$

$$\text{(A.9)}\qquad \frac{\partial^2\left[W_x^C + W_y^C\right]}{\partial (\tau^c)^2} = \frac{\partial M_x}{\partial \tau^c}, \qquad \frac{\partial^2\left[W_x^C + W_y^C\right]}{\partial (t^c)^2} = \frac{\partial Q_m}{\partial t^c}$$

Finally, from (2) and (3), one can derive that

$$\text{(A.10)}\qquad [t^D - s] = [\tau^N - \tau^c]\left(\frac{\partial M_x/\partial t}{\partial Q_m/\partial t}\right) + \left[M_x(\tau^c, t^D) - M_x(\tau^N, s)\right]\frac{\partial p^w/\partial t}{\partial Q_m/\partial t}$$

By Taylor series expansion,

$$\text{(A.11)}\qquad M_x(\tau^c, t^D) - M_x(\tau^N, s) = \frac{\partial M_x}{\partial t}[t^D - s] + \frac{\partial M_x}{\partial \tau}[\tau^c - \tau^N]$$

Substituting (A.11) into (A.10), one derives that

$$\text{(A.12)}\qquad [t^D - s] = \frac{\partial M_x/\partial t - (\partial M_x/\partial \tau)(\partial p^w/\partial t)}{\partial Q_m/\partial t - (\partial M_x/\partial t)(\partial p^w/\partial t)}[\tau^N - \tau^c]$$

Substituting the respective derivatives and (A.12) into (A.5) and (A.6), one derives that (7) is satisfied if

$$\text{(A.13)} \qquad \left|\frac{\partial Q_m/\partial t^c}{\partial M_x/\partial t^c}\right| \geq \left|\frac{\partial p^w/\partial t^c}{\partial p^w/\partial \tau^c}\right|$$

Since $|\partial Q/\partial t| > |\partial M/\partial t|$ and $|\partial p^w/\partial t| < |\partial p^w/\partial \tau|$, Equation (A.13) is always satisfied. ■

REFERENCES

Abrego, L., C. Perroni, J. Whalley, and R. Wigle, "Trade and Environment: Bargaining Outcomes from Linked Negotiations," *Review of International Economics* 9 (2001), 414–28.

Abreu, D., "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica* 56 (1988), 383–96.

Bagwell, K., and R. Staiger, "A Theory of Managed Trade," *American Economic Review* 80 (1990), 779–95.

———, and ———, "An Economic Theory of GATT," *American Economic Review* 89 (1999), 215–48.

Bernheim, D., and M. Whinston, "Multimarket Contact and Collusive Behavior," *RAND Journal of Economics* 21 (1990), 1–26.

Carraro, C., and D. Siniscalco, "Policy Coordination for Sustainability: Commitments, Transfers and Linked Negotiations," in I. Goldin and A. Winter, eds., *The Economics of Sustainable Development* (Cambridge: Cambridge University Press, 1995), 264–82.

———, and ———, "R&D Cooperation and the Stability of International Environmental Agreements," in C. Carraro, ed., *International Environmental Negotiations: Strategic Policy Issues* (Cheltenham: Edward Elgar, 1997), 71–95.

Cesar, H., and A. de Zeeuw, "Issue Linkage in Global Environmental Problems," in A. Xepapadeas, ed., *Economic Policy for the Environment and Natural Resources* (Cheltenham: Edward Elgar, 1996), 158–73.

Charnovitz, S., "Linking Topics in Treaties," *University of Pennsylvania Journal of International Economic Law* 19 (1998), 329–45.

Conconi, P., and C. Perroni, "Issue Linkage and Issue Tie-in in Multilateral Negotiations," *Journal of International Economics* 57 (2002), 423–47.

Copeland, B., "Strategic Interaction among Nations: Negotiable and Non-Negotiable Trade Barriers," *Canadian Journal of Economics* 23 (1990), 64–108.

Dixit, A., "Strategic Aspects of Trade Policy," in T. Bewley, ed., *Advances in Economic Theory: Fifth World Congress* (New York: Cambridge University Press, 1987), pp. 329–62.

Ederington, J., "International Coordination of Trade and Domestic Policies," *American Economic Review* 91 (2001), 1580–93.

Folmer, H., P. van Mouche, and S. Ragland, "Interconnected Games and International Environmental Problems," *Environmental and Resource Economics* 3 (August 1993), 313–35.

Friedman, J., "A Non-Cooperative Equilibrium for Supergames," *Review of Economic Studies* 28 (January 1971), 1–12.

Johnson, H., "Optimum Tariffs and Retaliation," *Review of Economic Studies* 21 (1953–54), 142–53.

Ludema, R., "Optimal International Trade Agreements and Dispute Settlement Procedures," *European Journal of Political Economy* 17 (2001) 355–76.

McCutcheon, B., "Do Meetings in Smoke-Filled Rooms Facilitate Collusion," *Journal of Political Economy* 105 (1997), 330–50.

Sebenius, J., "Negotiation Arithmetic: Adding and Subtracting Issues and Parties," *International Organization* 37 (1983), 281–316.

Spagnolo, G., "Issue Linkage, Delegation and International Policy Cooperation," Working Paper No. 49.96, Fondazione Eni Enrico Mattei, 1996.

———, "On Interdependent Supergames: Multimarket Contact, Concavity and Collusion," *Journal of Economic Theory* 89 (1999), 127–39.

van Damme, E., "Renegotiation-Proof Equilibria in Repeated Prisoner's Dilemma," *Journal of Economic Theory* 47 (1989), 206–17.