Domestic Policies, Hidden Protection and the GATT/WTO

Claustre Bajona
Ryerson University

Josh Ederington
University of Kentucky

October 2009

Abstract

As tariff barriers have fallen worldwide, regulation of domestic policy has become increasingly important in international trade agreements. This has led to the emergence of a theoretical literature addressing the integration of perfectly observable domestic policy into trade agreements. However, the assumption that domestic policy is perfectly observable is problematic since the interpretation and enforcement of domestic policy statues is often non-transparent. Thus, it may be difficult to determine whether lack of market access is due simply to random shocks or to the use of domestic policies as hidden trade barriers. In this paper, we model international coordination over trade and domestic policy when domestic policy is private information and thus can be used as a form of "hidden protection". We show that the optimal design of an efficient agreement depends greatly on whether domestic policy is observable or unobservable.

JEL Classification: F1, F13, K33, H2
KEYWORDS: trade policy, environmental standards, international agreements

1 Introduction

Under the auspices of the General Agreement on Tariffs and Trade (GATT), the international community has made great strides in lowering tariff barriers to trade. However, as tariff barriers have fallen, regulation of domestic policies, such as environmental regulations or labor standards, has become increasingly important in international trade agreements. A key concern is that, as countries enter into trade agreements that constrain their ability to use trade policy as a means of erecting barriers to trade, they will respond by using environmental or other domestic policies as a secondary means of protection (e.g., see Markusen (1975), Copeland (1990), Ederington (2001) and Bagwell and Staiger (2001)). Specifically, with respect to trade policy, the GATT is an instrument-based agreement, in that World Trade Organization (WTO) members negotiate over tariff concessions, and then are required by the agreement to respect the binding tariff ceilings that are a result of these negotiations. In contrast, the GATT is a rules-based agreement with respect to domestic policy and leaves discretion to WTO members in the setting of their domestic environmental and labor standards with the exception that the resulting standards must adhere to certain GATT rules (principally the national treatment provision in Article III of GATT). Thus, to the extent that countries can relax regulatory standards in import-competing sectors as a means of reducing trade flows, they can undermine commitments previously made with respect to trade policy. These concerns, that countries may manipulate domestic rules and regulations as a secondary trade barrier, has led to the emergence of a theoretical literature addressing the integration of domestic policy into trade agreements.

A key concern in this literature concerns the optimal design and structure of trade agreements when domestic policies are an (imperfect) substitute for trade policy as a means of influencing trade flows. However, this literature typically investigates international cooperation over domestic policy when the domestic policy choices of the each government is perfectly observable. We find this assumption problematic. Most basically, even when domestic policy statutes are transparent, the actual use and enforcement of such statutes is decidedly less so. Thus, it may be difficult to determine whether a reduction in foreign market access is due simply to random shocks or to the use of domestic policies as hidden trade barriers. In this paper, we model international coordination over trade and domestic policy when domestic policy is private information and show that several of the key lessons from information models fail to transfer fully to an imperfect-information setting.

Specifically, we first analyze the "national sovereignty" argument of Bagwell and Staiger (2001): that an efficient international trade agreement covering multiple policy instruments only need specify a minimum level of foreign market access (the actual policy mix in achieving that level of access can be chosen unilaterally). The "national sovereignty" argument is based on the fact that the key inefficiency is not the policy mix, but an inefficiently low level of foreign market access when policies are set unilaterally. As we show in this paper, the results of Bagwell and Staiger (2001) extend naturally to the case of self-enforcing agreements (i.e., agreements where cooperation is maintained by the threat of future punishment if cheating occurs) provided that both trade policy and domestic policy are perfectly observable. However, this is no longer the case when domestic policy is unobservable (i.e., the foreign country cannot distinguish between hidden protection and random fluctuations in import volume). Rather, we show that in addition to specifying a minimum level of market access, a well-designed international agreement should also establish a binding tariff ceiling as a means of forcing countries wishing to deviate from the agreement to use hidden domestic policy as the deviating policy. Since domestic policy is a second-best instrument for restricting market access, the corresponding incentive to deviate from the agreement is less and greater cooperation can be maintained.

Secondly we analyze the "linkage" argument of Ederington (2002) and Limao (2005): that, in the absence of transboundary non-pecuniary externalities, allowing cross-retaliation across policy instruments (e.g., threatening trade policy sanctions for deviations in domestic policy) is equivalent to having separate "unlinked" agreements which forbid such cross-retaliation. The "linkage" argument is based on the fact that, under conditions of perfect information, countries have a limited incentive to deviate from an agreement by using domestic policy (since it is a second-best means of influencing trade flows). Thus, the threat of threat of sanctions is unnecessary as a means of enforcing the agreement. However, such sanctions are also not costly since, with perfectly observable policies, the agreement can be structured so that deviations never occur and thus the threatened sanctions are never utilitized. In this paper we demonstrated that this equivalence is overturned when domestic policy is unobservable. Note that when domestic policy is unobservable there is a positive probability that punishment will be triggered when no cheating has taken place (as a low level of market access could be a signal that the foreign country is deviating from the agreement, or it could be simply due to unobservable market fluctuations). Thus, the potential for mistaken punishment implies that limitations should be placed on the severity of punishment that occurs. However, interestingly , we show that domestic policy unobservability actually increases the desirability for "linkage" since countries are more likely to use domestic policy as a means of hidden protection and thus the stronger threat of trade sanctions is not desirable as means of keeping these incentives in check.

Section 2 of the paper presents the basic model of trade and governmental policy. Section 3 considers the optimal design of an agreement when domestic policy is perfectly observable, while Section 4 considers the case of unobservable domestic policy. Section 5 discusses the applicability of the national sovereignty argument while Section 6 discusses linkage. Finally, Section 7 concludes.

2 The model

In this section we construct a standard general equilibrium model of trade between two countries, a home country (country 1) and a foreign country (country 2). The government of each country has access to two policy instruments to affect the volume of this trade: a trade instrument (tariff on the imported good) and a domestic policy instrument (tax/subsidy to production in the importing sector). Each country has the ability to produce three goods: a homogeneous good, $j = 0$, and two differentiated goods, $j = 1, 2$, all of which are traded. We assume that they produce the homogeneous good using identical technologies, but that the home country (country 1) has an absolute advantage in producing good 2 and the foreign country has an absolute advantage in producing good 1. Under these assumptions, good 1 is the import good of country 1 and good 2 is the import good of country 2.

In what follows we describe in detail the model economy. We assume that countries are completely symmetric and omit the country's superscript when not needed.[1]

2.1 Consumers

Each country is populated by a measure 1 of identical, infinitely-lived consumers who discount the future at a rate δ. Consumers derive utility from consumption of the three goods. We assume that preferences are quasi-linear in the homogeneous good and can be represented by the period utility function:

$$u(c) = c_0 + \frac{1}{z}[c_1 - \frac{c_1^2}{2}] + \frac{1}{z}[c_2 - \frac{c_2^2}{2}] \tag{1}$$

where c_j represents consumption of good j, and $z > 0$ is a parameter influencing the elasticity of demand. Consumers are endowed with ℓ units of labor which they supply inelastically.

2.2 Producers

2.2.1 Homogeneous good

The homogeneous good technology transforms labor inputs into product at a rate of 1-1. The production function can be written as,

$$y_0 = l_0 \tag{2}$$

Notice that under this production function, in equilibrium the economy's wage is equal to the price of the homogeneous good. In what follows we normalize the price of the homogeneous good to 1 and, thus, the equilibrium wage will also be equal to 1.

2.2.2 Differentiated goods

The differentiated goods are produced using labor and an industry-specific fixed factor in a Cobb-Douglas technology.[2] Without loss of generality, we normalize the amount of the fixed factor in each industry to be 1. We assume that the home country has an absolute advantage in producing good 2 and the foreign country has an absolute advantage in producing good 1. Specifically, the production function for the export good in each country is given by:

$$y_j = [2l_j]^{\frac{1}{2}} \tag{3}$$

[1]For simplicity, we assume that no international borrowing or lending is allowed and that there is no storage technology available. Under these assumptions consumers and producers face the same problem every period, as a function of government policies, and the model is static.

[2]These assumptions on preferences and technology greatly simplify the model, since they make policies in one country independent of policies in the other country. In particular, under these assumptions the world price of each differentiated good is independent of the policies affecting the other differentiated good.

where y_j represents production of good $j = 1, 2$. Likewise the production function for the import good in each country is given by:

$$y_j = l_j^{1/2} + \phi \tag{4}$$

Note that the importing sector of each country is affected by period productivity shocks. We assume that the shock, ϕ, takes on the value 0 with probability $1 - \alpha$ and the value $\theta > 0$ with probability α. We assume that the period value of ϕ is known to firms and consumers when making their period decisions and is fixed at a level to avoid reversals on the importing sector.[3]

2.3 Government

The government's objective is to choose policy instruments in order to maximize consumers' expected lifetime welfare. The policy instruments that are available to the government are trade policy (a tariff on the imported differentiated good), τ, and domestic policy (a subsidy on production of the imported differentiated good), t.

Every period, given government policies, consumers and firms choose their optimal behavior after productivity shocks are realized. In what follows, we denote as "static competitive equilibrium" the period solution of the consumers' and firms' problems as functions of the government policies.

2.4 Equilibrium

2.4.1 Competitive equilibrium of the static problem

In this subsection we characterize the competitive equilibrium of the static problem of this economy, *given the government policies*, and given the realization of the technology shocks. In the next subsection we solve for the optimal policy.

The consumers'objective is to maximize their expected lifetime utility. Given that there is no storage technology in the model, and international borrowing and lending is not allowed, consumers spend all their income in the period consumption goods and, therefore, they solve the following static problem every period:

$$\begin{aligned} \max \quad & \left\{ c_0 + \tfrac{1}{z}[c_1 - \tfrac{c_1^2}{2}] + \tfrac{1}{z}[c_2 - \tfrac{c_2^2}{2}] \right\} \\ \text{s.t.} \quad & c_0 + q_1 c_1 + q_2 c_2 = m \\ & c_0 \geq 0 \end{aligned} \tag{5}$$

where q_j represents the consumer's price of good j and m is the consumer's income. In this formulation of the problem we have normalized the price of the homogeneous good to 1.

From the first order conditions, assuming that the parameters are such that consumers consume the homogeneous good in equilibrium, we derive the expression for the demand of each good as

[3] Given equilibrium policies, this requires that $\theta < 1/(2 + 2z)$.

functions of prices and income, m:

$$\begin{aligned} c_1 &= 1 - zq_1 \\ c_2 &= 1 - zq_2 \\ c_0 &= m - (1 - zq_1)q_1 - (1 - zq_2)q_2 \end{aligned} \tag{6}$$

Firms in the differentiated sectors $j = 1, 2$ maximize period profits subject to the technology constraint:

$$\begin{aligned} \max \quad & \{p_j y_j - w l_j\} \\ & y_j = f(l_j) \end{aligned} \tag{7}$$

where p_j denotes the producer's price of the good and $f(l_j)$ denotes the production function of the good.

Profit maximization in the numeraire sector sets the wage rate in the economy at one. From the first order conditions of the above problem, we derive labor demanded, output and profits for the home country (country 1) as functions of prices (expressions for the foreign country are symmetrically defined):

$$l_1 = (\frac{p_1}{2})^2 \; and \; l_2 = p_2^2 \tag{8}$$

$$y_1 = \frac{p_1}{2} + \phi \; and \; y_2 = p_2 \tag{9}$$

$$\pi_1 = \frac{p_1^2}{4} + p_1\phi \; and \; \pi_2 = \frac{p_2^2}{2} \tag{10}$$

where π_j are the profits of industry $j = 1, 2$, and ϕ is the realization of the shock in country 1.

We assume that consumers own the firms in the economy. The government redistributes revenue from taxation and tariffs back to the consumers. Therefore, we can write the consumer's income as:

$$m^i = w^i l^i + \sum_j \pi_j + \tau^i M_i^i - t^i y_i^i \tag{11}$$

where M_i^i represents net imports of good i by country i.

The government in this economy imposes specific tariff rates on the imported good, τ, and gives specific subsidies to the domestic sector t. Given these policies, the relationship between consumer and producer prices in country i, which imports good i and exports good $-i$, is the following:

$$\begin{aligned} q_i^i &= p_i^w + \tau^i \quad , \quad p_i^i = p_i^w + \tau^i + t^i \\ q_{-i}^i &= p_{-i}^i = p_{-i}^w \end{aligned} \tag{12}$$

Definition 1 *A **competitive equilibrium of the static problem** of this economy is a sequence of functions of the government policies and technology shocks: consumer and producer decisions,* $\left\{c_j^i, y_j^i, l_j^i\right\}$ *and prices* $\left\{q_j^i, p_j^i, p_j^w, w^i\right\}$*, for* $i, j = 1, 2$ *such that:*
(i) given prices and income as defined in (11), $\left\{c_j^i\right\}$ *solve the consumers problem*
(ii) given prices, $\left\{y_j^i, l_j^i\right\}$ *solve the producers problem*
(iii) goods and labor markets clear, that is:

$$\begin{aligned} c_j^1 + c_j^2 &= y_j^1 + y_j^2 \quad j = 0, 1, 2 \\ l_0^i + l_1^i + l_2^i &= l^i \quad i = 1, 2 \end{aligned} \tag{13}$$

Notice that, since consumers face the same static problem every period, a competitive equilibrium for the economy is just a sequence of static competitive equilibria.

In what follows we derive analytical expressions for some of the variables in the model. From the market clearing conditions, one can solve for world prices of each good $j = 1, 2$ as a function of government policy and the productivity shock:

$$p_j^w = \frac{4 - t^j - 2\phi^j - \tau^j - 2z\tau^j}{3 + 4z} \tag{14}$$

Given these world prices, we can also determine import volume for each country $i = 1, 2$:

$$M_i^i = \frac{1 - t^i(1+z) - 2(1+z)\phi^i - \tau^i - z(3+2z)\tau^i}{3 + 4z} \tag{15}$$

The period welfare for each country, given domestic policy and shocks is given by the consumers' utility:

$$\hat{W}^i(t^1, \tau^1, t^2, \tau^2, \phi^1, \phi^2) = c_0^i + \frac{1}{z}\left(c_1^i - \frac{(c_1^i)^2}{2}\right) + \frac{1}{z}\left(c_2^i - \frac{(c_2^i)^2}{2}\right) \tag{16}$$

which, using the expressions for c_0^i and income, it becomes:

$$\begin{aligned} \hat{W}^i(t^1, \tau^1, t^2, \tau^2, \phi^1, \phi^2) = \quad & \ell^i + \pi_1^i + \pi_2^i + \tau^i M_i^i - t^i y_i^i - c_1^i q_1^i - c_2^i q_2^i \\ & + \tfrac{1}{z}\left(c_1^i - \tfrac{(c_1^i)^2}{2}\right) + \tfrac{1}{z}\left(c_2^i - \tfrac{(c_2^i)^2}{2}\right) \\ & \equiv \hat{W}_1^i(t^1, \tau^1, \phi^1) + \hat{W}_2^i(t^2, \tau^2, \phi^2) \end{aligned} \tag{17}$$

Notice that under our specifications, welfare in each country can be split into two components: one that depends only on the country's policies and realization of the shock, and another one that depends only on the other country's policy and realization of the shock.

2.4.2 Optimal policy

We assume that governments maximize expected national welfare (the aggregate utility of the representative consumers within the country). Given the distribution for the productivity shock, the expected welfare function can be written as:

$$\begin{aligned} W^i(t^1,\tau^1,t^2,\tau^2) = \; & \alpha^2 \hat{W}^i(t^1,\tau^1,t^2,\tau^2,\theta,\theta) \\ & +(1-\alpha)\alpha\left(\hat{W}^i(t^1,\tau^1,t^2,\tau^2,0,\theta)+\hat{W}^i(t^1,\tau^1,t^2,\tau^2,\theta,0)\right) \\ & +(1-\alpha)^2\hat{W}^i(t^1,\tau^1,t^2,\tau^2,0,0). \end{aligned} \tag{18}$$

2.4.3 Unilateral optimal Nash policies

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 the derivatives of $W^1(t^1,\tau^1,t^2,\tau^2)$ with respect to t^1 and τ^1, and solving out the first order conditions, the unilaterally optimal trade and domestic policies for the home country (t^D,τ^D) are given by:[4]

$$\begin{aligned} \tau^D(t) &= \tfrac{(1+2z)(1-2(1+z)\alpha\theta)-2(2+5z+3z^2)t}{2(2+9z+13z^2+6z^3)} \\ t^D(\tau) &= \tfrac{1-2(1+z)\alpha\theta-2(1+z)(2+3z)\tau}{4+z(11+8z)} \end{aligned} \tag{19}$$

In what follows, we eliminate the country superscripts for the domestic country (country 1) when it does not lead to confusion. So a policy or a variable without a country superscript is understood to refer to the home country.

Note from (19) that free-trade is not unilaterally optimal for the home country. Within this model, both countries are large and can influence the terms-of-trade with either their trade or domestic policies. Thus, there exists a unilateral incentive for each country to erect barriers to trade (with either tariffs or domestic taxes) as a means of pursuing terms-of-trade gains. From (19) one derives that the Nash equilibrium trade and domestic policies for each country will be defined by:

$$\begin{aligned} t^N &= 0 \\ \tau^N &= \tfrac{(1+2z)(1-2(1+z)\alpha\theta)}{2(2+9z+13z^2+6z^3)} \end{aligned} \tag{20}$$

Note from (20) that each country is imposing a positive import tariff in order to restrict trade while setting non-distortionary domestic taxes. This is the standard result of welfare analysis: in the presence of an international trade distortion (the terms-of-trade distortion), trade policy is the first-best policy choice.

The previous section established that countries have a unilateral incentive to erect trade barriers. However, while imposing barriers to trade may be unilaterally optimal, it is not optimal from a worldwide standpoint. Globally efficient trade and domestic policies will be set to maximize joint

[4] Best-response functions for the foreign country are defined symmetrically. Since markets for the two goods are independent and export policies are prohibited, these optimal policy choices are independent of foreign policy. The assumptions made about the functional forms of the model also ensure that the second-order conditions are satisfied and that a unique Nash equilibrium (in which the first-order conditions of (19) and the equivalent conditions for the foreign country hold simultaneously with positive trade volume) exists.

welfare ($W^1 + W^2$), and will serve as the natural goals toward which countries strive when they cooperate. In this paper, we focus on symmetric international agreements in which countries set common cooperative trade policies ($\tau^1 = \tau^2 = \tau^c$) and domestic policies ($t^1 = t^2 = t^c$), since common policies within the 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 expected welfare will be equivalent to maximization of a single country's expected welfare. Define $W^c(t^c, \tau^c)$ as the cooperative level of welfare (i.e., $W(\tau^1 = \tau^2 = \tau^c, t^1 = t^2 = t^c) = W^c(t^c, \tau^c)$). Taking derivatives of $W^c(t^c, \tau^c)$ with respect to t^c and τ^c one finds that globally efficient trade ($\overline{\tau}^c$) and domestic ($\overline{t}^c$) policies are given by free trade ($\overline{\tau}^c = 0$) and the non-distortionary domestic taxes ($\overline{t}^c = 0$). Intuitively, there is no reason for policy intervention driven by beggar-thy-neighbor, trade-restricting motivations in an efficient cooperative arrangement. 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.

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 direct enforcement, 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). Bagwell and Staiger (1990) and Dixit (1987) argue that countries can support lower trade barriers in a repeated game setting by threatening to retaliate against countries that deviate from the agreement. Thus, in these models, the threat of retaliation assists in maintaining lower trade barriers. In what follows we describe self-enforcing agreements for both a version of the model without uncertainty and then the model used in this paper.

3 Self-enforcing Agreements with Perfectly Observable Domestic Policy

In this section, we consider a situation with complete certainty: both trade and domestic policy are perfectly observable and the productivity levels of each country are fixed at $\phi^1 = \phi^2 = \theta$, that is, $\alpha = 1$. We focus on sub-game perfect trade strategies in which deviating from cooperative policies is punished by reversion to the static Nash-equilibrium tariffs:

Definition 2 *A self-enforcing agreement in this environment is characterized by common cooperative policies ($\tau^1 = \tau^2 = \tau^c$ and $t^1 = t^2 = t^c$) and a length of the punishment period $T \geq 0$ such that* *any defection from the established cooperative policies is punished by reversion to a punishment path for T periods.*

In what follows we consider the length of the punishment period T as exogenously given, and we derive formulas and results as functions of T. At the end of the section we discuss solving for the optimal length of the punishment period T. In order to simplify notation, we drop the dependence of T from the value functions and the cooperative policies.

For a given T, if each country were to set τ^c as their trade policy and t^c as their domestic policy, then punishment would never be triggered and:

$$V(\tau^c, t^c) \equiv \frac{1}{1-\delta} W(\tau^1 = \tau^2 = \tau^c, t^1 = t^2 = t^c) \tag{21}$$

where V denotes the discounted value of cooperating with the agreement and $\delta \in (0,1)$ represents the discount factor (the weight placed on future welfare).

If a country chooses to deviate from the agreement, it can do so by deviating in either trade or domestic policy only, or in both policies. If a country deviates, it will obtain the period welfare associated with the deviation and it will be reverted to the punishment phase for T periods. The present discounted value of each possible deviation is given by:

$$D_t(\tau^c, t^c) \equiv W_t^D + \delta V_t^P \tag{22}$$

$$D_\tau(\tau^c, t^c) \equiv W_\tau^D + \delta V_\tau^P \tag{23}$$

$$D_{\tau,t}(\tau^c, t^c) \equiv W_{\tau,t}^D + \delta V_{\tau,t}^P \tag{24}$$

where the subscripts indicate in which policy the country is deviating, $W_t^D \equiv W(\tau^1 = \tau^c, t^1 = t^D(\tau^c), \tau^2 = \tau^c, t^2 = t^c)$ is the welfare of unilaterally deviating in domestic policy, $W_\tau^D \equiv W(\tau^1 = \tau^D(t^c), t^1 = t^c, \tau^2 = \tau^c, t^2 = t^c)$ is the welfare of unilaterally deviating in trade policy, and $W_{\tau,t}^D \equiv W(\tau^1 = \tau^N, t^1 = t^N, \tau^2 = \tau^c, t^2 = t^c)$ is the welfare of unilaterally deviating to the full Nash equilibrium described in (20). The value functions V_i^P, $i = t, \tau, (\tau, t)$ represent the discounted value of entering into the punishment phase and are is equal to:

$$V_t^P = \beta^T W_t^P + \delta^T D_t \tag{25}$$

$$V_\tau^P = \beta^T W_\tau^P + \delta^T D_\tau \tag{26}$$

$$V_{\tau,t}^P = \beta^T W^P + \delta^T D_{\tau,t} \tag{27}$$

where $\beta^T = (1-\delta^T)/(1-\delta)$ is the overall discount factor that applies to the periods where the punishment is applied and W_i^P is the period welfare obtained during the punishment phase. In a linked agreement, punishment reverts to the full Nash equilibrium and, thus, $W_i^P = W^N \equiv W(\tau^1 =$

$\tau^2 = \tau^N, t^1 = t^2 = t^N$) for all $i = t, \tau, (\tau, t)$. In non-linked agreements, the deviating in one policy only will revert to Nash equilibrium in that policy only (see section 3.2.) Simplifying these expressions we obtain that the discounted values of deviating from the cooperative agreement as:

$$D_t(\tau^c, t^c) = \frac{1}{1-\delta^{T+1}} \left(W_t^D + \beta^T \delta W_t^P \right) \tag{28}$$

$$D_\tau(\tau^c, t^c) = \frac{1}{1-\delta^{T+1}} \left(W_\tau^D + \beta^T \delta W_\tau^P \right)) \tag{29}$$

$$D_{\tau,t}(\tau^c, t^c) = \frac{1}{1-\delta^{T+1}} \left(W_{\tau,t}^D + \beta^T \delta W^N \right) \tag{30}$$

For an agreement to be viable, it must be the case that neither country has an incentive to deviate from the agreement in either policy. Thus, an international agreement results in both countries jointly choosing trade and domestic policies to maximize the cooperative level of welfare, subject to a set of self-enforcement constraints (which entail balancing the gain to remaining in the agreement versus the gain to deviation). The problem that we need to solve is:

$$\begin{array}{ll} \max_{T,\tau^c,t^c} & V \\ s.t. & V \geq D_{\tau,t} \\ & V \geq D_t \\ & V \geq D_\tau \end{array} \tag{31}$$

In what follows, unless stated otherwise, we concentrate on linked agreements. In a linked agreement any deviation reverts to full Nash during the punishment phase, that is, $W_t^P = W_\tau^P = W^N$. Therefore, for any given T, since the period welfare of deviating in both policies is at least as high as the period welfare of deviating in only one policy, it is the case that $D_{\tau,t} \geq D_t$ and $D_{\tau,t} \geq D_\tau$. Therefore, the only constraint that binds in the maximization problem above is the first one: $V \geq D_{\tau,t}$. This implies that, within a linked agreement, countries solve the problem:

$$\begin{array}{ll} \max_{T,\tau^c,t^c} & V \\ s.t. & V \geq D_{\tau,t} \end{array} \tag{32}$$

The solution to the above maximization is a set of policies $\hat{\tau}^c, \hat{t}^c$ that we refer to as "most-cooperative policies", and a length of the punishment period T. The next subsection characterizes the solution to the problem described above.

3.1 Characterization of the optimal solution

3.1.1 Optimal length of the punishment period

Given that policies are perfectly observable, punishment is only triggered by deviating from the most-cooperative policies. There is no random deviation to the punishment phase. Therefore, the value of cooperating, V does not depend on T, and T affects only the present values of deviating.

Larger T lower the value of deviating in any given set of policies, reducing the incentives to deviate and allowing for greater degrees of cooperation. In a world where punishment cannot be triggered randomly, the optimal punishment period is $T = \infty$. The next lemma formalizes this result.

LEMMA 1 *In a self-enforcing agreement with perfectly observable domestic policy, the optimal punishment period is $T = \infty$.*

Proof:
See appendix.

3.1.2 Characterization of the most-cooperative policies

In this section we characterize the most-cooperative policies $(\hat{\tau}^c, \hat{t}^c)$ in the case where $T = \infty$. The below lemma shows that it is always optimal to set the domestic policy to zero $\hat{t}^c = 0$. This results is simply a reflection of the basic argument of Ederington (2001), that within a self-enforcing agreement covering both trade and domestic policy (and in the absence of a non-pecuniary transboundary externality), only cooperation in trade policy will be relaxed in order to satisfy the self-enforcement constraint. In addition, the proposition below gives the expression for the most-cooperative trade policy.

LEMMA 2 *Within a self-enforcing international agreement with perfectly observable domestic policy, the most-cooperative domestic policy is non-distortionary (i.e., $\hat{t}^c = 0$) and the most-cooperative trade policy is given by:*

$$\hat{\tau}^c = \frac{(4 + 6z - \delta(7 + 10z))\,(1 - 2(1 + z)\theta)}{2(2 + 5z + 3z^2)\,(4 + 6z - \delta(1 + 2z))} \tag{33}$$

if $\delta < (4 + 6z)/(7 + 10z)$, and by $\hat{\tau}^c = 0$ if $\delta \geq (4 + 6z)/(7 + 10z)$..

Proof:
See appendix.

The intuition behind the above result rests on first-best principles. The underlying reason countries want to defect from the international trade agreement is trade related (note that the international externality within the model is the terms-of-trade externality). Thus, allowing countries protection in trade policy, the most efficient means of affecting trade, will be the most efficient means of countering the incentive to deviate. Therefore, a cooperative agreement will result in setting domestic policy optimally (i.e., $\hat{t}^c = 0$) and lowering tariff barriers as far as the self-enforcement constraint allows. The above lemma shows that, if $\delta \geq (4 + 6z)/(7 + 10z)$ then the globally efficient policies ($\hat{t}^c = \hat{\tau}^c = 0$) are self-enforcing. If $\delta < (4 + 6z)/(7 + 10z)$ then $\hat{\tau}^c$ is given by (33), and

$\hat{t}^c = 0$. In the rest of the section, we extend to our framework several well-known results about the structure of optimal agreements given perfect information.

3.2 National Sovereignty

The concerns of many proponents of more fully incorporating environmental policy into trade agreements is that, as countries reciprocally increase trade flows through negotiated tariff concessions, they will attempt to substitute by degrading domestic standards so as to lower the production costs of domestic firms and thus reduce the level of market access provided to trading partners (i.e., the race-to-the-bottom problem). However, the ability to distort domestic standards as a secondary trade barrier was well understood by the drafters of the GATT and, as discussed in previous sections, these concerns manifested themselves in prohibitions against "discriminatory" standards as articulated in the national treatment rules of Article III of GATT. The key insight of Bagwell and Staiger (2001) is that these race-to-the-bottom concerns can also be addressed through "non-violation" complaints provided in Article XXIII of GATT. Specifically, if a WTO member can show that the market access commitments which it had previously negotiated are being offset by an unanticipated change in the environmental policies of another member country, then it has a right to seek redress even if the policy change was non-discriminatory and thus broke no explicit WTO rule. Thus, the right to bring nonviolation complaints can restrain foreign governments from using their domestic policies as secondary trade barriers. In this sense, the WTO can address race-to-the-bottom concerns even in the absence of explicit negotiations over the setting of domestic policy. Bagwell and Staiger (2001) show that an agreement that involves direct negotiations over tariff levels but allows countries to maintain autonomy over domestic policy would be inefficient. Specifically, it would result in governments subsequently distorting their domestic standards as a secondary means of reducing the market access achieved through tariff liberalization. However, they then consider the effect of adding the right to bring nonviolation complaints to the agreement. That is, they consider a two-stage negotiating game in which governments first use tariff negotiations to achieve an efficient level of market access (i.e., trade volume) and then are allowed sovereignty in the setting of policy provided that such policy does not undermine the market access commitments previously agreed upon. They show that such an agreement is efficient even though it does not require explicit negotiation over domestic policy. Intuitively, this is due to the fact that the underlying problem the agreements seek to correct is the level of market access (which is inefficiently low due to terms-of-trade considerations) and not the policy mix chosen by governments. Thus, once an agreement specifies the level of market access through negotiation, countries can be allowed sovereignty in choosing the efficient mix of trade and domestic policies to achieve that level of market access.

Thus, the "national sovereignty" result of Bagwell and Staiger (2001) argues that an efficient agreement does not need to specify a set of cooperative policies τ^c, t^c. Rather, full cooperation can be achieved simply by setting a minimum level of market access $\hat{M}$. It should be noted that Bagwell and Staiger (2001) consider the case where an international agreement can be externally enforced and thus enforcement constraints do not bind. Therefore, we must extend the "national sovereignty" result to the context of as self-enforcing agreement in which cooperation is sustained by the threat of future retaliation. An obvious analog is to consider the case where, rather than triggering punishment on deviating from a set of cooperative policies ($\tau \neq \tau^c$ or $t \neq t^c$), punishment is triggered when import volume falls below some minimum threshold (i.e., $M(\tau, t) < \hat{M}$). As we show in the following lemma, if domestic policy is observable, consistent with Bagwell and Staiger (2001), an import volume trigger can support the same degree of cooperation as a conventional agreement which specifies cooperative policies.

LEMMA 3 *Within a self-enforcing international agreement with perfectly observable domestic policy, the most-cooperative equilibrium can be obtained by an agreement that simply establishes a minimum level of market access (i.e., an import volume level) that each country must maintain and sets $T = \infty$.*

Proof:
Fix $T = \infty$, the optimal length of the punishment period for the most-cooperative equilibrium. Define $\hat{M}(\hat{\tau^c}, \hat{t^c}, \phi)$ as the most-cooperative import volume level (where $\hat{\tau^c}$ and $\hat{t^c}$ are the most-cooperative polices solved for in the previous subsection and ϕ is the productivity level). Let each country choose its trade and domestic policy unilaterally, subject to a minimum level of market access defined by $\hat{M}$. Assume that this level of market access is enforced by the threat of reversion to the Nash equilibrium for T number of periods (i.e., setting $M(t, \tau, \phi) < \hat{M}$ will trigger punishment). Note that, since most-cooperative policies are self-enforcing, $\hat{M}$ is self-enforcing as well. In this case, an agreement which sets $\hat{M}$ will result in the the home country choosing t and τ to maximize $W(t, \tau, t^2, \tau^2)$ subject to $M(t, \tau, \phi) \geq \hat{M}(\hat{\tau^c}, \hat{t^c}, \phi)$). From the first-order conditions of this constrained maximization, one derives that :

$$\frac{\partial W(\tau, t, \tau^2, t^2)/\partial \tau}{\partial W(\tau, t, \tau^2, t^2)/\partial t} = \frac{\partial M_1(\tau, t)/\partial \tau}{\partial M_1(\tau, t)/\partial t} \tag{34}$$

Taking the above derivatives, one derives that (34) reduces to:

$$\frac{-2t(1+z)(2+3z) + (1+2z)A}{A - t(4 + z(11 + 8z))} = 1 + 2z, \tag{35}$$

where $A = 1 - 2(1+z)\phi - 2(1+z)(2+3z)\tau$.

First, note that (35) is satisfied by setting non-distortionary domestic policy (i.e., $t = \hat{t}^c = 0$). Second, note that, since the most-cooperative import volume is greater than the unilaterally optimal import volume, the country will raise its tariff until the import volume constraint binds (i.e., $M_1(t, \tau, \phi) = \hat{M}(\hat{\tau^c}, \hat{t^c}, \phi)$). Since $t = \hat{t}^c = 0$, this implies that $\tau = \hat{\tau}^c$). Q.E.D.

It should be noted first that the above lemma is a slightly restated version of the result of Bagwell and Staiger (2001). The results in that paper were referring specifically to globally efficient policies (unconstrained by limited enforcement power). In contrast, this paper is considering self-enforcing agreements. However, as can be seen from the above lemma, the results of Bagwell and Staiger (2001) naturally extend to the case of self-enforcing agreements. Second, note that the import volume trigger specified by an efficient agreement, $\hat{M}(\hat{\tau^c}, \hat{t^c}, \phi)$ is a function of the import volume shock, ϕ. Thus, to the extent that their exist random market shocks, the minimum level of market access specified by the agreement will vary from period to period in line with those random shocks to productivity. Such an adjustment requires perfect information about market conditions in each country.

3.3 Linkage

A key question in the literature on self-enforcing trade agreements covering multiple policy instruments is whether such an agreement should be configured so that cheating on any part of it would trigger a costly retaliatory episode covering all parts of the agreement, or should retaliation be confined to the provisions where the cheating took place. For example, should one enforce obligations with respect to domestic policy (e.g., environmental or labor standards) within an international agreement with the threat of the suspension of trade concessions. A recent literature (e.g., Ederington (2002), Limao (2005)) has emerged that attempts to provide theoretical insight into this question.[5] The basic difference between the linked and non-linked agreements is that the linked agreement is configured so that cheating on any part of the agreement triggers retaliation in all parts of the agreement, while the non-linked agreement confines the retaliation to that part of the agreement where the cheating took place. Specifically, in both linked and non-linked agreements, deviation in both trade and domestic policy will trigger reversion to the Nash equilibrium in both policies from the linked agreement ($D^L_{\tau,t}$) and the non-linked agreement ($D^{NL}_{\tau,t}$) is given by (30). Thus, the self-enforcement constraint with respect to a deviation in both policies (i.e., $V \geq D_{\tau,t}$) is identical for the two types of agreements. However, in contrast to the linked agreement, deviation in a single policy from the cooperative equilibrium of a non-linked agreement triggers reversion to the Nash equilibrium in only that policy. Define $W^N_\tau \equiv W(\tau^1 = \tau^D(t^c), t^1 = t^c, \tau^2 = \tau^D(t^c), t^2 = t^c)$ as the period welfare from a reversion to the Nash equilibrium in trade policy. In that case, the present discounted value of deviating from the cooperative agreement only in trade policy are, using (29):

$$D^L_\tau = \frac{1}{1-\delta^{T+1}} \left(W^D_\tau + \beta^T \delta W^N \right) \tag{36}$$

[5] Also see Spagnolo (1996) on the potential benefits of linkage over multiple policy issues.

for the linked agreement, and

$$D_{\tau}^{NL} = \frac{1}{1-\delta^{T+1}}\left(W_{\tau}^{D} + \beta^{T}\delta W_{\tau}^{N}\right) \tag{37}$$

for the non-linked agreement.

Likewise, defining $W_t^N \equiv W(\tau^1 = \tau^c, t^1 = t^D(\tau^c), \tau^2 = \tau^c, t^2 = t^D(\tau^c))$ as the Nash equilibrium on domestic policy, the present discounted value of deviating from the agreement in domestic policy alone is given by, using (30):

$$D_{t}^{L} = \frac{1}{1-\delta^{T+1}}\left(W_{t}^{D} + \beta^{T}\delta W^{N}\right) \tag{38}$$

for the linked agreement, and

$$D_{t}^{NL} = \frac{1}{1-\delta^{T+1}}\left(W_{t}^{D} + \beta^{T}\delta W_{t}^{N}\right) \tag{39}$$

for the non-linked agreement.

Notice that lemma 1 applies to non-linked agreements as well, and the optimal length of the punishment period is also $T = \infty$. Therefore, with both a linked and a non-linked agreement countries choose $T = \infty$, and τ^c and t^c that maximize V subject to the self-enforcement constraints that:

$$V \geq D_{\tau,t}\ ,\ V \geq D_{\tau}\ ,\ V \geq D_{t} \tag{40}$$

It should be apparent from the above discussion that the linked agreement threatens a tougher retaliatory episode for when a country deviates in only a single policy. It is well known in the game theoretic literature, that given perfect information the punishment phase is never triggered in equilibrium and such stronger punishment is always weakly preferred (see Abreu (1988).) Thus, the previous theoretical literature has focused not on the optimality of linkage, but rather the conditions under which it is strictly preferable to non-linkage. With respect to our analysis, it is instructive to consider whether, under the conditions of the model introduced in this paper, linkage is in fact necessary. Indeed, as we derive below, a non-linked agreement can support the same degree of cooperation as the linked agreement (i.e., the linked and non-linked agreements are functionally equivalent):

LEMMA 4 *Within a self-enforcing agreement with perfectly observable domestic policy, linked and non-linked agreements can support the same degree of cooperation (i.e., $\hat{\tau}^c, \hat{t}^c$).*

Proof:

Using lemmas 1 and 2, at the optimal solution of both types of agreements satisfies $T = \infty$ and $\hat{t}^c = 0$. We, thus, evaluate the self-enforcement constraints at the point where $\hat{t}^c = 0$. The self-enforcement constraint with respect to a deviation in both policies (i.e., $V \geq D_{\tau,t}$) is identical for *the two types of agreement, and the lowest cooperative tariff* $\hat{\tau_L}^c$ *that satisfies this constraint is given by (33) in lemma 3. The non-linked agreement faces two additional constraints that* $V \geq D_\tau^{NL}$ *and* $V \geq D_t^{NL}$. *Given* $\hat{t}^c = 0$, *it is direct to derive that the first additional constraint is binding for* $\hat{\tau}^c$ *given by (33). Regarding the second constraint, if we evaluate* $V - D_t^{NL}$ *for* $T = \infty$ *at the* $\hat{\tau_L}^c$ *we obtain:*

$$V - D_t^{NL} = \frac{z(3+4z)^2\delta^3\left(1-2(1+z)\theta\right)^2}{(2+3z)\left(4+z(11+8z)\right)^2(1-\delta)\left(4+6z-\delta(1+2z)\right)} \tag{41}$$

which is strictly positive. Therefore, the pair $\hat{\tau_L}^c, \hat{t}^c = 0$ *is feasible in the non-linked agreement. The optimal level of cooperation is, thus, at least as good as this one. Combining this result with Lemma 2 we obtain that both types of agreements can support the same degree of cooperation. Q.E.D.*

The above equivalence result is primarily derived from our assumption that there are no cross-border (non-pecuniary) externalities. In this situation, our framework fails to satisfy the conditions, specified in Limao (2005) of supermodularity in trade and environmental policy, for linkage to be strictly preferred.

4 Self-Enforcing Agreements with Unobservable Domestic Policy

In the above section, we considered optimal cooperation over trade and domestic policies given a situation of perfect information and full certainty. However, it seems natural to relax the assumption of perfect information, especially with respect to domestic policies and other non-tariff trade barriers. Indeed, almost all of the previous literature on negotiations over tariff and non-tariff trade barriers (e.g., Copeland (1990), Riezman (1991) and Hungerford (1991)) have assumed that non-tariff barriers were unobservable. As these papers stress, non-tariff trade barriers (especially domestic policies like labor and environmental standards) are not as transparent as tariffs, and thus are both more likely to be used by deviating countries (as hidden protection) and are more likely to be the cause of disagreements and trade wars across countries.

In this section, we adopt the view that domestic policies are less transparent and hence harder to monitor than trade policies. To formalize this view, we assume that the distribution of the productivity shock ϕ is non-degenerate (that is, $\alpha < 1$) and that countries cannot observe either the domestic policies of other countries or the realization of the random variable ϕ.[6] In contrast, the trade policy decisions of each country remain perfectly observable. In addition, uncertainty is

[6] A key difference between our model and that of Riezman (1991) is that he assumed that protection in general is not observable, since unobservable domestic policies can be constructed which are perfect substitutes for a tariff. In our framework, unobservable domestic policy is an imperfect substitute for the tariff. This can be justified by assuming that disguising protection from foreign observation places constraints on the abilities of governments to replicate the efficiency of tariffs as a form of protection.

introduced into the model by assuming that the productivity shock is only realized after governments make their policy decisions.

The non-observability of domestic policy raises the question of how cooperation can be maintained within an agreement if deviation from cooperative policies cannot be observed. The basic model (adopted from Green and Porter (1984) and Riezman (1991)) is one where cooperation is maintained by threatening retaliation if import volume falls below some trigger level. As noted by Riezman (1991), this trigger strategy has the advantage of being both simple and corresponding to actual and proposed policy measures. However, uncertainty within the model means that countries cannot always distinguish between hidden protection and random fluctuations in import volume. In the following sections, we analyze the respective costs and benefits to linking issues in the presence of this type of uncertainty. As in section 3, we derive formulas and results for the case where the length of the punishment period T is exogenously given. We discuss the case where T is optimally chosen in the last subsection.

4.1 Imperfect Information and Trigger Strategies

As mentioned previously, trade policy remains observable and thus punishment can be triggered by any observed deviation from a cooperative level of trade policy, τ^c. For maintaining cooperation in domestic policy, we follow Green and Porter (1984) by using trigger strategies where strategies are conditioned on observables and, as in Riezman (1991), use import volume as our trigger. Thus, if import volume falls below some predetermined critical level $\overline{M}$, countries will take that as evidence that cheating on the agreement has occurred and trigger the punishment phase.

We assume a self-enforcing agreement specifies a length of the punishment period T, symmetric "most-cooperative" policies to be played in each period, τ^c and t^c, and an import volume, $\overline{M}$, that triggers punishment. Following our analysis for the certainty case, we derive formulas and results as a function of the length of the punishment phase T and discuss optimality with respect to T in a separate subsection. For any given T, note that, it contrast to the case of perfect information and certainty, our import volume trigger can not be made conditional on the (unobservable) random shock. Thus, if the home country sets policy such that if $M(\tau, t, \phi = \theta) \geq \overline{M}$, where $M(\ )$ denotes the import function, then punishment is never triggered by the home country. In contrast, if policy is set so that $M(\tau, t, \phi = 0) < \overline{M}$ then punishment is triggered with probability one. Finally, if $M(\tau, t, \phi = 0) \geq \overline{M} > M(\tau, t, \phi = \theta)$ then punishment is triggered with probability between zero and one.

Fix the length of the punishment period T. For any a trigger strategy with cooperative policies (τ^c, t^c) and import level $\overline{M}$, we derive expressions for the discounted present value of cooperating and deviating from the domestic policy. We first need to introduce some notation. Define the probability of a country's triggering punishment when cooperating and deviating from the cooperative agreement, respectively, as:

$$\begin{aligned} \eta^C &= \Pr\{M(\tau^c, t^c, \phi) < \overline{M}\} \\ \eta^D &= \Pr\{M(\tau^c, t^D, \phi) < \overline{M}\} \end{aligned} \tag{42}$$

where t^D is the optimal domestic policy when cooperating in trade policy. Define $\gamma^C = (1-\eta^C)^2$ as the probability of *not* triggering punishment when both countries are collaborating, and $\gamma^D = (1-\eta^D)(1-\eta^C)$ as the probability of *not* triggering punishment when deviating from the cooperative domestic policy (while the other country cooperates). Notice that these probabilities depend on both the cooperative policies and the import level. Assume that the foreign country plays the cooperative policies. The expected value of playing the cooperative policies (τ^c, t^c) for the home country is then:

$$V(\tau^c, t^c; \overline{M}) = \frac{1}{1-\delta\gamma^C}\left(W(\tau^c, t^c) + \delta(1-\gamma^C)V^P\right) \tag{43}$$

where V^P is the present value of the punishment phase due to randomly triggered punishment (notice that no country has deviated from cooperative policies). Let W^P define the welfare of a period where punishment is implemented (this is either reversion to full Nash in a linked agreement, or reversion to Nash equilibrium in the deviating policy in a non-linked agreement). Then given that punishment lasts for T periods, V^P becomes:

$$V^P = \beta^T W^P + \delta^T V \tag{44}$$

Similarly, the present value of deviating in one policy only and of deviating in both policies can be written as:

$$D_t(\tau^c, t^c; \overline{M}) = \frac{1}{1-\delta\gamma^D}\left(W_t^D(\tau^c, t^D(\tau^c)) + \delta(1-\gamma^D)V_t^P\right) \tag{45}$$

$$D_\tau(\tau^c, t^c; \overline{M}) = W_\tau^D(\tau^D(t^c), t^c) + \delta V_\tau^P \tag{46}$$

$$D_{\tau,t}(\tau^c, t^c; \overline{M}) = W_{\tau,t}^D(\tau^N, t^N) + \delta V_{\tau,t}^P \tag{47}$$

where the present value of punishment in each of the cases is given by:

$$V_t^P = \beta^T W_t^P + \delta^T D_t \tag{48}$$

$$V_\tau^P = \beta^T W_\tau^P + \delta^T D_\tau \tag{49}$$

$$V_{\tau,t}^P = \beta^T W^N + \delta^T D_{\tau,t} \tag{50}$$

In the latter expressions W_t^P and W_τ^P represent the period welfare when punishment for deviating respectively in domestic and trade policy only is implemented. As in the previous case, this is either the period welfare in the full Nash equilibrium for a linked agreement or welfare in the Nash equilibrium in domestic policy or trade policy in a non-linked agreement. The punishment phase lasts for T periods in both types of agreements. Notice that, given that deviating in trade policy triggers punishment with probability one, in a linked agreement, deviating in trade policy is always dominated by deviating in both policies.

Simplifying these expressions we obtain:

$$V = \frac{1}{A^C}\left(W + \beta^T \delta(1-\gamma^C)W^P\right) \tag{51}$$

$$D_t = \frac{1}{A^D}\left(W_t^D + \beta^T \delta(1-\gamma^D)W_t^P\right) \tag{52}$$

$$D_\tau = \frac{1}{1-\delta^{T+1}}\left(W_\tau^D + \beta^T \delta W_\tau^P\right) \tag{53}$$

$$D_{\tau,t} = \frac{1}{1-\delta^{T+1}}\left(W_{\tau,t}^D + \beta^T \delta W^N\right) \tag{54}$$

where A^C and A^D are constants and are defined as $A^C = 1 - \delta\gamma^C - (1-\gamma^C)\delta^{T+1}$ and $A^D = 1 - \delta\gamma^D - (1-\gamma^D)\delta^{T+1}$.

We are interested in solving for the cooperative policies and import volume trigger that maximize the expected discounted value of welfare. That is, the problem that the governments face is to find an optimal length of the punishment period T, policies (τ^c, t^c), and trigger $\overline{M}$ that solve the maximization problem:

$$\begin{array}{ll} \max_{T,\tau^c,t^c,\overline{M}} & V \\ s.t. & V \geq D_{\tau,t} \\ & V \geq D_t \\ & V \geq D_\tau \end{array} \tag{55}$$

The following lemma simplifies the problem above by showing that given a length of the punishment phase T, for any given import volume trigger $\overline{M}$ and cooperative trade policy τ^c, there are only three possible symmetric equilibrium self-enforcing values for the domestic policy. This property is due to the fact that we have a discrete distribution for the productivity shock in which it can take only two values.

LEMMA 5 *Fix the length of the punishment period $T \geq 0$. Given a cooperative trade policy τ^c and an import volume trigger, $\overline{M}$, in a symmetric equilibrium a self-enforcing domestic policy t^c can take only three values:*

- *(1) $t^c = t^D$* *No Cooperation*
- *(2) t^c s.t. $M(\tau^c, t^c, \phi = \theta) = \overline{M}$* *Weak Trigger*
- *(3) t^c s.t. $M(\tau^c, t^c, \phi = 0) = \overline{M}$* *Strong Trigger*

Proof: Proof by contradiction: assume that t^c is such that $M(\tau^c, t^c, \phi = 0) < \overline{M}$. Then punishment next period is triggered with probability one and thus each country will play Nash policies (i.e.,

$t^c = t^D$). Next, assume that t^c is such that $M(\tau^c, t^c, \phi = \theta) > \overline{M}$. Then either country can make a small deviation in t from this equilibrium without triggering punishment. Unless $t^c = t^D$, such a deviation will be profitable and thus $M(\tau^c, t^c, \phi = \theta) > \overline{M}$ cannot be an equilibrium. Finally, assume t^c such that $M(\tau^c, t^c, \phi = 0) > \overline{M} > M(\tau^c, t^c, \phi = \theta)$. In this case, punishment is triggered with probability α, and either country can make a small deviation in t from this equilibrium without increasing the probability of punishment. Unless $t^c = t^D$, such a unilateral deviation will be profitable, and thus this cannot be an equilibrium. Q.E.D.

The above Lemma suggests that, for any given T, it may be possible for countries to maintain some degree of cooperation (i.e., a $t^c < t^D$) given the use of the import trigger strategy. We refer to one possible equilibrium path as a "Weak Trigger" since it reflects cooperative policies (τ^c, t^c) and a trigger level $(\overline{M})$ such that punishment is never triggered if both countries cooperate. We refer to another possible equilibrium path as a "Strong Trigger" since it reflects cooperative policies (τ^c, t^c) and a trigger level $(\overline{M})$ such that punishment is randomly triggered on the realization of the low import volume shock. In the following sections, we show that which of these strategies emerges in equilibrium depends on parameter values and we give examples in which the equilibrium presents a strong trigger and examples in which it presents a weak trigger.

4.2 Strong Trigger Strategies

In this section, we consider the case where the cooperative policies (τ^c, t^c) and trigger level $(\overline{M})$ are such that punishment is randomly triggered with positive probability even if countries cooperate. This is a strong trigger since any attempt to deviate from cooperative policies (in either trade or domestic policy) by reducing market access will trigger punishment with probability one.

Given T, each country will set domestic policy in each period to maximize the expected value of lifetime utility subject to the constraint that when import volume falls below $\overline{M}$ the punishment phase is triggered. In this case, the probability of a given country triggering punishment when collaborating is $\eta^C = \alpha$, and the probability of triggering punishment when deviating in domestic policy is $\eta^D = 1$. Notice that deviation in trade policy always triggers punishment with probability one. Therefore, the probability of punishment not being triggered when both countries collaborate is $\gamma^C = (1-\alpha)^2$ and the probability of not triggering punishment when one of the countries deviates, given that the other plays the cooperative policies, is $\gamma^D = 0$. That is, any deviation from the cooperative policies triggers punishment with probability one. The expected values of cooperating and deviating become:

$$V^s = \frac{1}{A^C}\left(W + \beta^T \delta\alpha(2-\alpha)W_t^P\right) \tag{56}$$

$$D_t^s = \frac{1}{1-\delta^{T+1}}\left(W_t^D + \beta^T \delta W_t^P\right) \tag{57}$$

$$D_\tau^s = \frac{1}{1-\delta^{T+1}}\left(W_\tau^D + \beta^T \delta W_\tau^P\right) \tag{58}$$

$$D_{\tau,t}^s = \frac{1}{1-\delta^{T+1}}\left(W_{\tau,t}^D + \beta^T \delta W^N\right) \tag{59}$$

where $A^C = 1 - \delta(1-\alpha)^2 - \alpha(2-\alpha)\delta^{T+1}$. Notice that in a linked agreement, given that deviation triggers punishment with probability one, and punishment reverts to the full Nash equilibrium ($W_t^P = W_\tau^P = W^N$), if a country is to deviate, it will do so in both policies, since $W_t^D \leq W_{\tau,t}^D$ and $W_\tau^D \leq W_{\tau,t}^D$.

It is direct to derive from the equations above that (τ^c, t^c) can be supported as a self-enforcing equilibrium with a strong trigger (i.e., $V^s \geq D_{\tau,t}^s$) if:

$$W(\tau^c, t^c) \geq \frac{1}{1-\delta^{T+1}}\left(A^c W_{\tau,t}^D + \beta^T \delta\left(1-\delta(1-\alpha)^2\right) W^N\right) \tag{60}$$

First, note from (60), that an increase in α reduces the degree of cooperation that can be maintained. Intuitively, this is due to the fact that if α is high, then cooperation is likely to be randomly abandoned and thus deviation is profitable. Second, note that with a strong trigger agreement, this is the only self-enforcement constraint that must be satisfied. Thus, for a given T, the situation of strong triggers is similar to the perfect certainty case of Section 3 in that any deviation triggers automatic punishment, and thus countries that choose to cheat will deviate to the Nash equilibrium. The characterization of the optimal policy is also similar to the case of perfect certainty, and is described briefly below. The two situations differ, however, in an important issue: in a strong trigger punishment may be triggered randomly: even if both countries cooperate, when the productivity shock in the import industry in one of the countries is high, punishment will be triggered. Even if both countries cooperate, reversal to the punishment phase occurs with some probability, and the discounted value of cooperation is a function of the length of the punishment period T. Therefore, setting $T = \infty$ is not necessarily optimal in this case. In what follows we characterize the optimal solution under a strong trigger. Unless stated otherwise, we assume that the agreements are linked.

4.2.1 Characterization of the optimal solution

In this section we characterize the optimal length of the punishment period T and the most-cooperative policies $(\hat{\tau}^c, \hat{t}^c)$ for a strong trigger. In the proposition below we show that, independently of the value of T, it is always optimal in a strong trigger to set the domestic policy to zero $\hat{t}^c = 0$. In addition, we derive the for the most-cooperative trade policy and the optimal length of the punishment period as functions of the parameters of the model.

PROPOSITION 1 *Within a self-enforcing international linked agreement with strong triggers and length of the punishment period given by T, the most-cooperative domestic policy is non-distortionary* *(i.e., $\hat{t}^c = 0$) and the optimal length of the punishment period $\hat{T}$ and the most-cooperative trade policy $\hat{\tau}^c$ are such that:*

$$\delta^{\hat{T}} = \frac{-4 - 6z + \delta(7 + b + 2z(5 + b))}{(3 + b + 2z(2 + b))\delta} \tag{61}$$

$$\hat{\tau}^c = \frac{b(1 - 2(1 + z)\alpha\theta)}{(1 + z)(3 + b + 2z(2 + b))} \tag{62}$$

if $\delta \geq (4 + 6z)/\,(7 + b_1 + 2z(5 + b)) \equiv \delta^{\min}$, where $b = (2 - \alpha)\alpha$, and

$$\hat{T} = \infty \tag{63}$$

$$\hat{\tau^c} = \tau^N \frac{4 + 6z - \delta(7 + 10z)(1 - b)}{4 + 6z - \delta(1 + 2z)(1 - b)} \tag{64}$$

if $\delta < \delta^{\min}$.

Proof:

In order to prove this theorem, note that the proof that $\hat{t}^c = 0$ follows that of lemma 2. Thus, we set $t^c = 0$ and for any given length of the punishment period T we compute the trade policy $\tau^c(T)$ that solves $V^T = D^T_{\tau,t}$. We then evaluate the discounted present value of cooperation at at these policies, $V^T(tau^c(T))$, and optimize over T. See the appendix for details and specific expressions. Q. E. D.

Notice that for all $0 < \delta < 1$, the optimal trade policy $\hat{\tau}^c$ satisfies $0 < \hat{\tau}^c < \tau^N$. Therefore, an optimal strong trigger always allows some cooperation but it can never achieve full cooperation. This result is clearly in contrast with the case with certainty, where full cooperation was achieved for some values of the parameters. The intuition for the result is simple and highlights the importance of triggering random punishment: in order to achieve full cooperation, "too much" punishment should be inflicted when punishment is triggered randomly. Utility maximizing agents stay away from such situations by choosing a shorter length of the punishment period and allowing only for partial cooperation.

4.3 Weak Trigger Strategies

In this section, we take the case where the cooperative policies (τ^c, t^c) and trigger level $\overline{M}$ are such that punishment is never triggered if both countries cooperate. This is a weak trigger since it allows deviations from cooperative policies to (possibly) go unpunished since sufficiently small deviations from t^c will not lower import volume sufficiently to trigger punishment in the event of a high import volume shock. In particular, a small deviation from the cooperative policy. Setting t such that $M(\tau^c, t, \phi = 0) \geq \overline{M} > M(\tau^c, t, \phi = \theta)$ will only trigger punishment with probability α, that is, on the realization of the low import volume shock. Let $\bar{t}^D = t^c + 2\theta$. This policy satisfies$M(\tau^c, \bar{t}^D, \phi = 0) = \overline{M}$. For any domestic policy $t \leq \bar{t}^D$, punishment is triggered only with probability α. Since unilateral welfare is monotonically increasing in t for $t \leq t^D(\tau^c)$ (the unilateral Nash deviation in domestic policy), the home country will choose the largest $t \in [0, t^D(\tau^c)]$ such that $M(\tau^c, t, \phi = 0) \geq \overline{M}$. Therefore, if the home country is to deviate in domestic policy only, it chooses $\hat{t}^D = \min(\bar{t}^D, t^D(\tau^c))$. That is, if the Nash deviation, $t^D(\tau^c)$, is not too large (that is, it triggers punishment only with probability α), then the country chooses to deviate to the Nash equilibrium in domestic policy. If the Nash deviation is too large (that is, it triggers punishment with probability one), then the country chooses the largest possible deviation that triggers punishment only with probability α: $\bar{t}^D$. The next lemma derives conditions on the parameters under which the Nash deviation is too large.

LEMMA 6 *In the model described above, fix T and let (τ^c, t^c) be the most cooperative policies that are to be achieved with a weak trigger strategy. Then, the deviation to the Nash equilibrium in domestic policy triggers punishment with probability α if and only if:*

$$\theta \geq \frac{1 - 2(1+z)(2+3z)\tau^c - Bt^c}{2B + 2(1+z)\alpha}, \tag{65}$$

where $B = 4 + z(11 + 8z)$. Therefore, if this condition is met, $\hat{t}^D = t^D(\tau^c)$. In any other case, $\hat{t}^D = \bar{t}^D$.

Proof: The result follows from direct comparison of the definitions of $t^D(\tau^c)$ in (19) and $\bar{t}^D = t^c + 2\theta$. Q.E.D.

Define $\hat{W}_t^D = W(\tau^1 = \tau^2 = \tau^c, t^1 = \hat{t}^D, t^2 = t^c)$. The expected discounted value of a small deviation (setting $t = \hat{t}^D$) is given by:

$$\hat{D}_t^w = \hat{W}_t^D + \delta \left[(1-\alpha)\hat{D}_t^w + \alpha V_t^P \right] \tag{66}$$

Simplifying this equation, we obtain:

$$\hat{D}_t^w = \frac{1}{A^D} \left(\hat{W}_t^D + \beta^T \delta \alpha W_t^P \right) \tag{67}$$

where A^D is now $A^D = 1 - \delta(1-\alpha) - \alpha\delta^{T+1}$.

If countries cooperate, punishment is triggered with probability zero in a weak trigger strategy. The expected discounted value of cooperating, then becomes:

$$V^w = \frac{W}{1-\delta} \tag{68}$$

Deviating in trade policy only or in both policies triggers punishment with probability one. Therefore, the expected discounted values of deviating in trade policy only and in both policies are, respectively:

$$D_\tau^w = \frac{1}{1-\delta^{T+1}} \left(W_\tau^D + \beta^T \delta W_\tau^P \right) \tag{69}$$

and

$$D^w_{\tau,t} = \frac{1}{1-\delta^{T+1}}\left(W^D_{\tau,t} + \beta^T \delta W^N\right) \tag{70}$$

Notice that these last two equations are the same as the corresponding equations for the strong trigger (equations 58 and 59). Therefore, the only differences between a strong and a weak trigger are in the expected discounted value of cooperating and of deviating in trade policy. Furthermore, a large deviation from the cooperative equilibrium in domestic policy (such that $M(\tau, t, \theta = 0) \leq \overline{M}$) or any deviation in trade policy will trigger punishment with probability one. In a linked agreement, if a country were to make such a deviation, it would deviate to the unilaterally optimal Nash policies (τ^N, t^N). With a weak trigger strategy, for (τ^c, t^c) to be self-enforcing the discounted value of cooperation needs to be higher than the value of deviating to full Nash equilibrium, and also higher than a small (hidden) deviation in domestic policy. That is:

$$V^w \geq D^w_{\tau,t} \tag{71}$$

and

$$V^w \geq \hat{D}^w_t \tag{72}$$

That is, using the expressions for the value functions, this implies:

$$W \geq \frac{1-\delta}{1-\delta^{T+1}}\left[W^D_{\tau,t} + \beta^T \delta W^N\right] \tag{73}$$

and

$$W \geq \frac{1-\delta}{A^D}\left[\hat{W}^D_t + \beta^T \delta \alpha W_N\right] \tag{74}$$

Notice that the first condition is the same as the binding enforcement constraint under perfect information and certainty.

In solving for the most cooperative policies under a weak trigger, it is key to determine which of the two equations above binds. We are interested in situations where the second condition is binding. Comparing the right hand sides of the equations above, this occurs as long as:

$$\hat{W}^D_t - \frac{A^D}{1-\delta^{T+1}} W^D_{\tau,t} \geq \beta^T \delta \frac{(1-\alpha)(1-\delta)}{(1-\delta^{T+1})A^D} W^N \tag{75}$$

We characterize the solution of a weak trigger strategy for each possible pair of parameter values (α, δ) for the values of θ such that condition (65) is satisfied. For parameter values for which condition (65), the behavior is quantitatively similar, but it is more complicated to solve for, since the limiting areas change with θ.

4.3.1 Characterization of solution

As in the certainty case, the present discounted value of cooperation V does not depend on the length of the punishment period and, therefore, $T = \infty$ is optimal. Given this, their exists two probabilities: (1) that parameter values are such that $V \geq D_{\tau,t}$ is the binding condition in the self-enforcing agreement and (2) that $V \geq D_t$ is the binding condition. The next proposition characterizes the set of parameters for which each constraint is binding, and solves for the optimal policies in each case. It is stated for values of θ for which condition (65) is satisfied. Figure 2 illustrates this characterization.

PROPOSITION 2 *Within a self-enforcing international linked agreement with weak triggers the optimal punishment period is $T = \infty$. Let $z = 1$, assume that condition (65) is satisfied. Define:*

$$\delta_1 = \frac{10}{17} \tag{76}$$

$$\delta_2 = \frac{200}{200 + 483\alpha} \tag{77}$$

$$\delta_3 = \frac{1610}{2877 - 3017\alpha + 724\sqrt{5\alpha - 5\alpha^2}} \tag{78}$$

For $20/69 < \alpha \leq 1$, $V \geq D_{\tau,t}$ is the binding condition. It is the case that:

- *If $\delta \geq \delta_1$, then $\hat{\tau}^c = \hat{t}^c = 0$ and the globally efficient policies are optimal*
- *If $\delta < \delta_1$, then $\hat{t}^c = 0$ and $\hat{\tau}^c = (10 - 17\delta)(1 - 4\alpha\theta)/(200 - 60\delta)$*

For $0 \leq \alpha \leq 20/69$, $V \geq D_t$ is the binding condition. It is the case that:

- *If $0 < \delta \leq \delta_3$, then $\hat{t}^c = t^N$ and $\hat{\tau}^c = \tau^N$ and no cooperation can be achieved*
- *If $\delta_3 < \delta < \delta_2$, then $\hat{t}^c = 7\hat{\tau}^c$ and $\hat{\tau}^c = (3620(1-\delta) + A)(1-4\alpha\theta)/\,(3620(181(1-\delta) + 161\alpha\delta))$, where $A = 7\sqrt{8326\alpha\delta(49 + 49\alpha + 69\alpha\delta)}$*
- *If $\delta_2 \leq \delta < 1$, then $\hat{t}^c = \hat{\tau}^c = 0$ and full cooperation is attained.*

Proof: The proof that $T = \infty$ follows that of lemma 2.

Second, note that the most-cooperative domestic policy takes the value $\hat{t}^c = 0$ if the binding condition is $V \geq D_{\tau,t}$ (this follows from the same proof as lemma 1). For the case when the binding condition is $V \geq D_t$ take the Lagrangian of the maximization with respect to t^c and τ^c and following the calculations in the proof to lemma 1, one derives the following condition:

$$\frac{\partial W^C/\partial\tau^c}{\partial W^C/\partial t^c} = \frac{\partial W_t^D((\tau^1 = \tau^2 = \tau^c, t^1 = \bar{t}^D, t^2 = t^c)/\partial\tau^c}{\partial W_t^D(\tau^1 = \tau^2 = \tau^c, t^1 = \bar{t}^D, t^2 = t^c)/\partial t^c} \tag{79}$$

Taking derivatives of the cooperative and deviating levels of welfare with respect to cooperative trade and domestic policies (and substituting in for $\bar{t}^D$) one derives that:

$$\frac{-2t^c - 6\tau^c}{-3t^c - 2\tau^c} = \frac{-14t^c - 40\phi - 42\tau^c}{-21t^c - 46\phi - 14\tau^c} \tag{80}$$

Thus, by (80), (79) is satisfied if $t^c = 7\tau^c$.

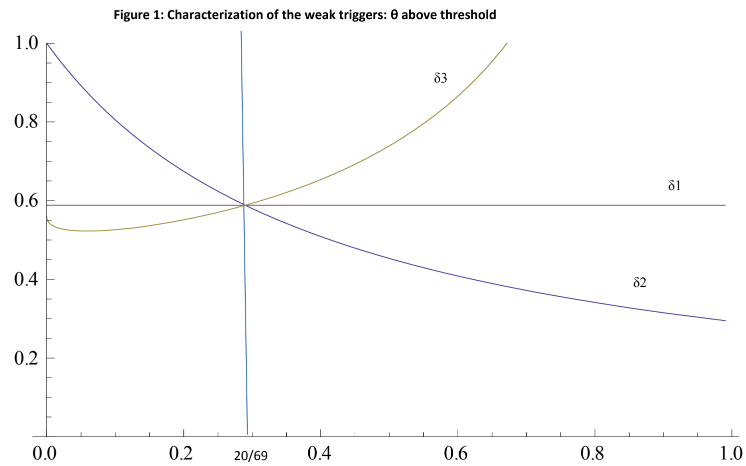

Figure 1: Characterization of the weak triggers

To finish the proof , we assume that each condition is binding and solve for the optimal $\hat{\tau}^c$ given the corresponding $\hat{t}^c$ from the previous proposition. We then compute the other constraint at the optimal values and find conditions under which the constraint is satisfied. Q. E. D.

Note that the above proposition implies that, under certain parameters, the binding constraint is $V \geq D_t$. In this case $\hat{t}^c \neq 0$ if $\hat{\tau}^c \neq 0$. That is, the domestic policy is distortionary unless the globally efficient policies $\tau^c = t^c = 0$ are self-enforcing. This result suggests that the results of Ederington (2001) no longer apply in a world of asymmetric information and uncertainty. The intuition behind allowing for distortions in domestic policy is direct. The hidden aspect of domestic policy makes it more attractive as a means of deviating from an international agreement (since one can distort one's private domestic policy to restrict trade without necessarily triggering foreign punishment). When domestic policy is more likely to be used as a means of cheating on the agreement, an efficient agreement will not require full cooperation in domestic policy. Rather, cooperation in domestic policy will be relaxed so as to ensure that the agreement is self-enforcing. This result is in direct contrast to the case where domestic policy is perfectly observable, in which case any efficient self-enforcing agreement will require non-distortionary domestic policy.

5 National Sovereignty with Imperfect Information

In this section we analyze the national sovereignty result of Bagwell and Staiger (2001) under conditions of imperfect information. Recall that Bagwell and Staiger (2001) argues that an efficient agreement does not need to specify a set of cooperative policies τ^c, t^c. Rather, full cooperation can be achieved simply by setting a minimum level of market access $\hat{M}$. We show that, in general, this result need not hold in the model with non-observable domestic policies.

5.1 Strong triggers

The analysis of the national sovereignty result for strong triggers is quite simple. Since any deviation triggers punishment with probability one, equilibrium under a strong trigger behaves much like the certainty case, and the national sovereignty result holds. The following proposition states this fact more formally:

PROPOSITION 3 *Within a self-enforcing international agreement with strong trigger strategies, the most-cooperative equilibrium can be obtained by an agreement that simply establishes a minimum level of market access (i.e., an import volume level) that each country must maintain.*

Proof:
Set the length of the punishment period to be equal to its optimal value $\hat{T}$. The proof then follows that of Lemma 2.

5.2 Weak triggers

With weak trigger strategies the "national sovereignty" argument of Bagwell and Staiger (2001) no longer applies in a world of uncertainty and imperfect information. Specifically, for some parameter values, maximal cooperation can no longer be achieved simply by specifying a minimum level of market access. Rather, an efficient agreement will specify not only a level of market access (i.e., the import volume trigger, $\overline{M}$), but also a binding tariff ceiling (i.e., $\overline{\tau}$). We illustrate this result using the example in figure 2.

Assume $z = 1$, and $\theta < 1/(46 + 4\alpha)$ (the result is qualitatively identical if this condition is not satisfied). Figure 2 shows the set of parameters (α, δ) for which the first best set of policies $\tau^c = t^c = 0$ can be achieved with a weak trigger strategy. In order to contradict the national sovereignty argument, we need to find a set of parameters for which full cooperation is attained with a weak trigger strategy when both a tariff ceiling τ^c and a minimum level of imports $\overline{M}$ are specified, but it cannot be achieved by specifying only a minimum level of imports (we call the latter a "B-S equilibrium").

The main difference between our equilibrium and the B-S equilibrium is that, since the B-S equilibrium only specifies a minimum level of imports, countries can make small deviations in trade policy and still trigger punishment only with probability α. Let the B-S minimum level of imports be a weak trigger. That is:

$$\overline{M} \equiv M(\tau^c = t^c = 0, \phi = \theta) = \frac{1-4\theta}{7} \tag{81}$$

The maximum tariff that triggers punishment with probability α, assuming no deviation in domestic policy, $\overline{\tau^D}$, solves the equation:

$$M(t^c = 0, \tau, \phi = 0) = \overline{M} \tag{82}$$

which is satisfied for $\overline{\tau^D} = 2\theta/3$. Notice that under our assumption on parameters, this tariff is smaller than the Nash tariff (for cases where it would be bigger, the deviation would be to Nash equilibrium in trade policy and it would work in a similar way). Let $D_\tau^{BS,w}$ be the value of deviating in trade policy only in the B-S equilibrium. Then a pair of policies are self-enforced if $V^W \geq D_\tau^{BS,w}$, which in our example it is satisfied for pairs of parameters (α, δ) such that:

$$\delta \geq \frac{800\theta\,(3-4(5+3\alpha)\theta)}{-504\alpha^2\theta + 800(3-20\theta)\theta + 1008\alpha^3\theta^2 + \alpha(63-9600\theta^2)} \equiv \delta_\tau^{BS,w}(\alpha) \tag{83}$$

Comparing this function with the one obtained in (77), we can show that $\delta_\tau^{BS,w} \geq \delta_{t2}^w$, with equality only when $\alpha = 0$ and $\delta = 1$. Therefore, there is a set of parameters for which a weak trigger in the B-S equilibrium cannot sustain full cooperation, whereas it can be sustained in our set up. This set is illustrated in figure 2.

As the above example illustrates, specifying a binding tariff ceiling can enlarge the scope of cooperation within the agreement. Thus, as we argue in the following proposition, having a binding tariff ceiling is weakly preferred (and in some cases strictly preferred) to simply specifying a minimum level of market access:

PROPOSITION 4 *Within a self-enforcing international agreement with weak trigger strategies, having an agreement that specifies a minimum level of market access ($\overline{M}$) and a binding tariff ceiling ($\overline{\tau}$) is weakly preferred and in some cases strictly preferred to an agreement that simply establishes a minimum level of market access.*

Proof:

First, assume an agreement that specifies both a binding tariff ceiling ($\overline{\tau}$) and minimum level of market access ($\overline{M}$) such that either $\tau > \overline{\tau}$ or $M(\tau, t, \phi) < \overline{M}$ will trigger reversion to the punishment phase. In this case, the self-enforcement constraints are given by (71) and (72).

Next, assume an agreement that specifies only an import volume trigger such that $M(\tau, t, \phi) < \overline{M}$ will trigger reversion to the punishment phase. In this case, small deviations in either *trade policy*

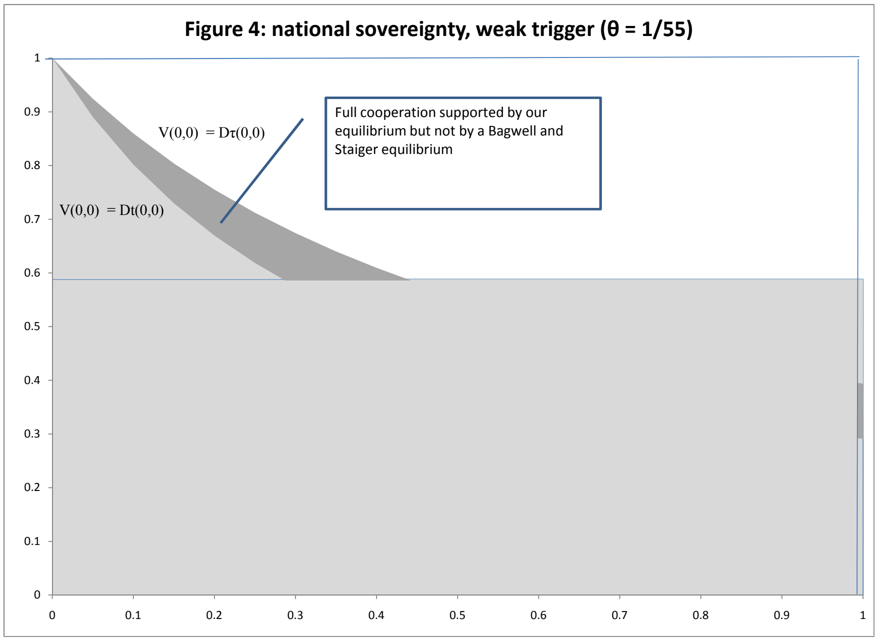

Figure 2: National Sovereignty, weak trigger

or domestic policy (such that $M(\tau, t, \phi = 0) \geq \overline{M} > M(\tau, t, \phi = \theta)$*) will only trigger punishment with probability* α*. Thus, in addition to (71) and (72), there is an additional constraint that no country has an incentive to make a small deviation in trade policy:*

$$V^w \geq \hat{D}_\tau^w \tag{84}$$

where $\hat{D}_\tau^w$ *represents the expected present value of a small deviation in trade policy that only triggers punishment with probability* α*. Given the presence of this additional constraint, an agreement that specifies only an import volume trigger cannot generate a higher degree of cooperation (i.e., a greater* V^w*) and, as we previously showed by example, in some cases results in a lower degree of cooperation. Q.E.D.*

This proposition shows that, in general, an efficient agreement should specify both a minimum level of market access ($\overline{M}$) and a binding tariff ceiling ($\overline{\tau}$). What is noteworthy about this result is, as is pointed out by Bagwell and Staiger (2001), this can be interpreted as exactly what the articles of GATT require. Specifically, tariff commitments are established by negotiated bound tariffs, while market access commitments are established by the non-violation complaints of Article XXIII.

In a situation of perfect information and certainty Bagwell and Staiger (2001) argue that binding a countries tariffs can impede the attainment of globally efficient outcomes by deterring countries from making changes to their domestic policies that would increase access to their markets. Thus, Bagwell and Staiger (2001) argue for providing greater sovereignty to countries in choosing their policy mix than is currently provided under GATT rules. Basically, they argue that countries should be allowed to raise tariffs above bound levels without triggering punishment (provided they maintain the specified level of imports). In contrast, we argue that current GATT restrictions on trade policy (i.e., bound tariffs) can be justified as a means of enforcing greater cooperation in a world of uncertainty about a country's domestic policies. Intuitively, this is because domestic policy is only a second-best means of restricting market access. Thus, since binding tariffs force countries to use less efficient domestic policy as a means of cheating on the agreement, it also reduces the incentives countries have to deviate from the agreement.

6 Linkage with Imperfect Information

Finally, we consider the potential benefits of linkage (i.e., allowing cross-retaliation over trade and domestic policy) in the presence of imperfect information. Note that, in contrast to the perfect certainty case, the presence of imperfect information can result in an equilibrium characterized by the random triggering of disputes and punishment (on the realization of bad market shocks). This suggests a potentially positive role for forbidding policy linkage as a means of reducing the loss to such random disputes. This possibility is considered in the next couple of sections.

6.1 Strong trigger

The main difference between a self-enforcing agreement with perfect certainty and a self-enforcing agreement with strong trigger strategies is the addition of a random probability that reversion to the punishment phase will be triggered. An important question this raises is whether a non-linked agreement would be preferable as a means of reducing the loss to randomly generated punishment.

As in the linked agreement, the expected, discounted value of cooperation in a non-linked agreement is given by (56). However, in a linked agreement any deviation (such that either tariffs are above their binding levels or import volume is below its binding level) triggers reversion to the full Nash equilibrium and thus $W_t^P = W^N$. In contrast, in a non-linked agreement a deviation in solely domestic policy (so that the import volume falls below its binding level, but tariffs remain at co-operative levels) triggers reversion to the Nash equilibrium in domestic policy alone (i.e., only the agreement covering domestic policy is abandoned). Thus, in the non-linked agreement $W_t^P = W_t^N$). Since reversion to the full Nash equilibrium is more severe (i.e., $W_t^N \geq W^N$ for any $\hat{\tau}^c, \hat{t}^c$), the discounted value of cooperating in a non-linked agreement is actually higher than the discounted value of cooperating in a linked agreement (i.e., $V^{NL}(\tau^c, t^c) > V^L(\tau^c, t^c)$). These calculations offer a potential explanation for why, in the presence of uncertainty, a linked agreement covering multiple policy instruments may be suboptimal. Of course the benefit of the stronger punishment of linkage is that it typically allows for a greater degree of cooperation to be maintained (i.e., it allows lower τ^c and t^c to be supported as a self-enforcing equilibrium). However, as we show in the Proposition below, in a strong trigger strategy, the increased punishment of policy linkage is unnecessary, and thus non-linkage is strictly preferred. For simplicity we prove the proposition for the case $z = 1$, but the lemma should apply for any value of z.

PROPOSITION 5 *Assume $z = 1$. Within a self-enforcing international agreement with unobservable domestic policy but strong trigger strategies, forbidding cross-retaliation across policy instruments is optimal (i.e., a non-linked agreement is strictly preferred to a linked agreement).*

Proof:

Set T to be the optimal length of the punishment period for the linked agreement. Define V^{NL} and V^L as the discounted value of cooperation in a non-linked and a linked agreement, respectively. Define $D^{NL}_{\tau,t}$ ($D^L_{\tau,t}$)as the discounted value of deviating in both policies from a non-linked (linked) agreement. V^{NL} is given by (56) where $W^P_t = W^N_t$. If a country deviates in both policies, such that $\tau > \tau^c$ and $M(\tau, t, \phi) < \overline{M}$, this triggers reversion to the full Nash equilibrium. The expected value of such a deviation is given by (59). Thus, for a given τ^c and t^c, $V^{NL} > V^L$ and $D^{NL}_{\tau,t} = D^L_{\tau,t}$, and the self-enforcement constraint for multiple policy deviations is less binding in the non-linked agreement.

Alternatively, a country could deviate in solely domestic policy, such that $\tau = \tau^c$ and $M(\tau^c, t, \phi) < \overline{M}$, thus triggering reversion to the Nash equilibrium in domestic policy alone. The expected value of such a deviation in a non-linked agreement is given by, following (57):

$$D^{NL}_t \equiv \frac{1}{A^D}\left(W^D_t + \beta^T\delta(1-\gamma^D)W^N_t\right) \tag{85}$$

Let $(\hat{\tau}^c, \hat{t}^c = 0)$ be the most-cooperative policies for the linked agreement. Then $V^L(\hat{\tau^c}, 0) = D_{\tau,t}(\hat{\tau^c}, 0)$. Given that $W^N_t > W^N$, we have that $V^{NL}(\hat{\tau^c}, 0) > V^L(\hat{\tau^c}, 0)$, and this constraint is not binding for the non-linked agreement (recall that the value of deviating in both policies is the same for the linked and non-linked agreement). With $z_1 = z_2 = 1$ and $T = \infty$, the most cooperative trade policy $\hat{\tau}^c$ is:

$$\hat{\tau}^c = \frac{1}{20A}\Big(10 - 17(1-\alpha)^2\delta + (7 + 17(-2+\alpha)\alpha\delta^{T+2}\Big(1 - 4\alpha\theta) \tag{86}$$

where $A = 10 - 3(1-\alpha)^2\delta\left(-7 + 3(-2+\alpha)\alpha\right)\delta^{T+1}$. Replacing $\hat{\tau}^c$ with this value in $V^{NL}(\hat{\tau^c}, 0)$ and $D^{NL}_t(\hat{\tau^c}, 0)$, we obtain that $V^{NL} \geq D^{NL}_t$ if and only if:

$$\frac{49(1-\alpha)^6\delta^3(1-\delta^T)^3(1-4\alpha\theta)^2}{2645A^2(1-(1-\alpha)^2\delta + (-2+\alpha)\alpha\delta^{T+1})} \geq 0, \tag{87}$$

which is clearly satisfied. Intuitively, if domestic policy is a less efficient means of restricting market access, and punishment is triggered with probability one for any deviation, countries have less *incentive to deviate in domestic policy. Thus, the self-enforcement constraint for a single issue deviation is also less binding in the non-linked agreement. Since none of the enforcement constraints for the non-linked agreement are violated at* $(\hat{\tau^c}, 0)$*, the expected discounted value at the most cooperative policies for the linked agreement is greater or equal to* $V^{NL}(\hat{\tau^c}, 0)$ *and therefore, given that* $V^{NL} > V^L$*, the non-linked agreement is strictly preferred Q.E.D.*

A common criticism of linkage is that conflicts with respect to the domestic policy portion of the agreement (e.g., disputes of environmental regulations) may undermine international trade agreements. As we have shown, such concerns can be formalized by a model in which domestic policy is unobservable. Any agreement covering unobservable policies must rely on some type of trigger strategy to induce cooperation. However, given random, unobservable shocks, such agreements carry the risk that punishment will be periodically triggered even when no deviation has taken place.[7] Therefore, countries will be concerned with minimizing the losses to reverting to the punishment phase. What we argue in the above proposition is that using trade policy sanctions to punish deviations from the import volume trigger is unnecessarily severe punishment (i.e., it is not necessary to deter deviations in domestic policy). Thus, linking trade and domestic policy agreements, which minimizes the losses from random punishment episodes, will be strictly preferred.

6.2 Weak triggers

Now consider the case of weak triggers where punishment is never triggered as long as countries set cooperative policies. One might expect the stronger punishment of linkage to be weakly preferred in an agreement with weak trigger strategies since, even in the presence of unobservable domestic policy, the punishment phase is never triggered. As we show in the following proposition, this is exactly the case:

PROPOSITION 6 *Within a self-enforcing international agreement with unobservable domestic policy but weak trigger strategies, allowing cross-retaliation across policy instruments is optimal (i.e., a linked agreement is weakly preferred and in some cases strictly preferred to a non-linked agreement).*

Proof:

Consider the case where (72) binds. In a linked agreement, a small deviation in domestic policy will trigger reversion to the Nash equilibrium in both policies (such that $W_t^P = W^N$*). In contrast, in a non-linked agreement the identical deviation in domestic policy will only trigger reversion to the Nash equilibrium in domestic policy (such that* $W_t^P = W_t^N$*). The fact that linkage is weakly optimal*

[7] It should be noted that carrying out such punishment is not a "mistake", but is necessary to make the punishment threat credible and thus deter deviations.

simply follows from the fact that punishment through reversion to the Nash equilibrium in both policies is greater than punishment in a single policy instrument ($W_t^N \geq W^N$) and thus linkage relaxes the constraint (72). Notice that since a weak trigger strategy triggers punishment with probability zero when both countries cooperate, $V^{NL} = V^L$. The case where the linked agreement is strictly preferred can be proved by example.

However, in contrast to the perfect certainty case, the assumption of unobservable domestic policy increases the incentive to deviate in domestic policy alone, and thus breaks the equivalence between linked and non-linked agreements:

Let $z = 1$ and $\theta < 1/(46+4\alpha)$ (the result is qualitatively identical if this condition is not satisfied). Set $T = \infty$, its optimal value for both types of agreements. Figure 5 plots the set of parameters (α, δ) for which full cooperation can be achieved with a linked and non-linked agreement. In particular, the figure adds the restriction of cooperation in domestic policy in the non-linked agreement to figure 2B. This condition is given by:

$$\delta \geq \frac{2116\theta\,(1-(23+4\alpha)\theta)}{-168\alpha^2\theta + 2116(1-23\theta)\theta + 336\alpha^3\theta^2 + \alpha(21-8464\theta^2)} \equiv \delta_{t2}^{wNL}(\alpha) \tag{88}$$

We observe that the set of parameters for which the linked agreement can sustain full cooperation is larger. Therefore, for parameters in the set A, full cooperation can be achieved under the linked agreement only. Since with a weak trigger strategy $V^{NL} = V^L$, the linked agreement is, thus, strictly better for this set of parameters.

Intuitively, although linkage and non-linkage are equivalent with perfect information, the non-transparency of domestic policy makes it a more attractive instrument with which to deviate. Thus, while the self-enforcement constraints for single policy deviations are non-binding in a world of perfect certainty (thus, resulting in the equivalence of linked and non-linked agreements), they can bind in a world of imperfect information. Such a result is ironic in a sense as many trade economists are skeptical about the wisdom of allowing cross-retaliation, fearing that the use of trade policy sanctions as a means of enforcing cooperation over a host of issues might undermine the advances that have been achieved in reducing trade barriers. Specifically, they fear that allowing cross-retaliation could result in disputes covering domestic policy to spill over into the trade arena. Such concerns are moot in a world of perfect information as trade disputes never actually occur in equilibrium (and linkage and non-linkage are functionally equivalent). However, the above proposition suggests that the presence of uncertainty and imperfect information could actually increase the desirability of cross-retaliation and the case for policy linkage.

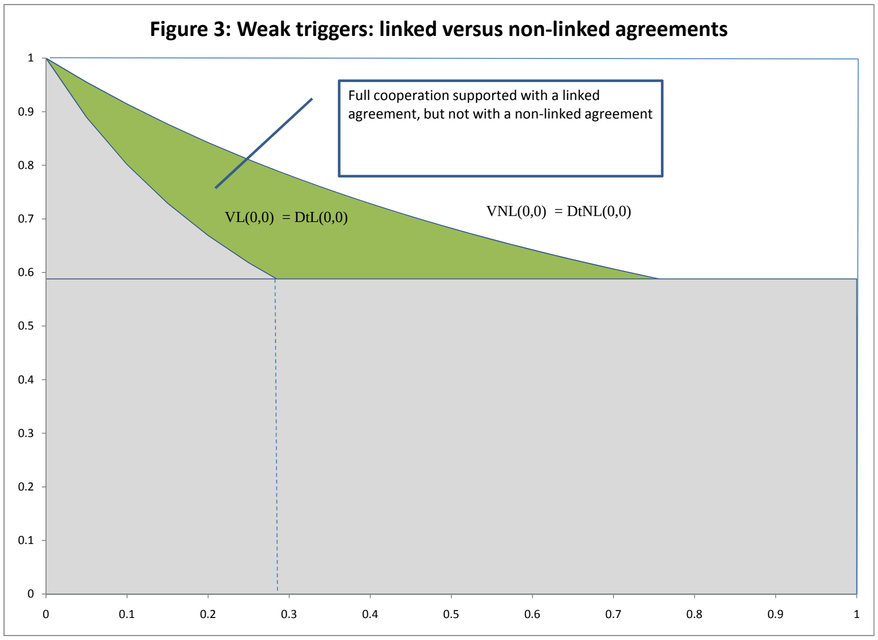

Figure 3: Linked versus non-linked agreements, weak trigger

7 Conclusion

A key question in recent trade negotiations is how to integrate domestic polices (such as environmental regulations and labor standards) into conventional trade agreements. In this paper, we consider the efficient design of such an international trade agreement when trade policy is observable but domestic policy is unobservable. We consider this a relevant situation since, while trade barriers might be readily apparent to foreign countries, the actual enforcement of many purely domestic regulations is not. The basic conclusion of this paper is that the observability of domestic policy is crucial as many of the main rules that govern the efficient treatment of domestic policy in a world of perfect information do not extend to a situation of asymmetric information.

8 Appendix

8.1 Proof of lemma 1

From equations (27) and (28)-(30) it is immediate to see that the value of cooperating, V does not depend on T, and the values of deviating in any set of policies, D_i,$i = t, \tau, (\tau, t)$, are decreasing functions of T. Let (τ^c, t^c) be a feasible policies under a given length of the punishment period $T < \infty$. Then it is the case that $V^T(\tau^c, t^c) \geq D^T_{\tau,t}(\tau^c, t^c)$, where the superscript T indicates the length of the punishment period. Since for any pair of policies $V^T = V^\infty$, and $D^T_{\tau,t} > D^\infty_{\tau,t}$, it is also the case that $V^\infty(\tau^c, t^c) \geq D^\infty_{\tau,t}(\tau^c, t^c)$, and the policies (τ^c, t^c) are feasible for $T = \infty$. Therefore, the set of feasible policies is the largest at $T = \infty$ and since the value of cooperation is independent of T, it will reach its highest value for $T = \infty$. Q. E. D.

8.2 Proof of lemma 2

Taking the Lagrangian of the above maximization with respect to t^c and τ^c one derives the following first-order conditions for both the linked and non-linked agreements:

$$\begin{array}{c} \frac{\partial V}{\partial t} - \lambda[\frac{\partial V}{\partial t} - \frac{\partial D_{t,\tau}}{\partial t}] = 0 \\ \frac{\partial V}{\partial \tau} - \lambda[\frac{\partial V}{\partial \tau} - \frac{\partial D_{t,\tau}}{\partial \tau}] = 0 \end{array} \tag{89}$$

It is direct to calculate, from the definitions of V and $D_{t,\tau}$ that:

$$\begin{array}{c} \frac{\partial V}{\partial t^c} = \frac{1}{1-\delta}\frac{\partial W^c}{\partial t^c} \quad and \quad \frac{\partial V}{\partial \tau^c} = \frac{1}{1-\delta}\frac{\partial W^c}{\partial \tau^c} \\ \frac{\partial D_{t,\tau}}{\partial t^c} = \frac{\partial W^D_{t,\tau}}{\partial t^c} \quad and \quad \frac{\partial D_{t,\tau}}{\partial \tau^c} = \frac{\partial W^D_{t,\tau}}{\partial \tau^c} \end{array} \tag{90}$$

Substituting (90) into (89) and solving out for the Lagrange multiplier, one derives the condition for an interior maximization of welfare in either cooperative agreement:

$$\frac{\partial W^c/\partial \tau^c}{\partial W^c/\partial t^c} = \frac{\partial W^D_{t,\tau}/\partial \tau^c}{\partial W^D_{t,\tau}/\partial t^c} \tag{91}$$

Taking derivatives of the cooperative and deviating levels of welfare with respect to cooperative trade and domestic policies, one derives that:

$$\frac{\partial W^C/\partial \tau^c}{\partial W^C/\partial t^c} = \frac{-2t^c - 6\tau^c}{-3t^c - 2\tau^c} \tag{92}$$

$$\frac{\partial W^D_{t,\tau}/\partial \tau^c}{\partial W^D_{t,\tau}/\partial t^c} = \frac{-3 + 6t^c + 12\phi + 18\tau^c}{-1 + 2t^c + 4\phi + 6\tau^c} \tag{93}$$

By (93) and (92), (91) is satisfied if $t^c = 0$, and thus the most-cooperative domestic policy is non-distortionary.

Finally, we compute the trade policy at which the self-enforcement constraint $V \geq D_{\tau,t}$ evaluated at $\hat{t}^c = 0$ is binding. We obtain $\hat{\tau}^c$ as described in (33). Notice that $\hat{\tau}^c > 0$ if and only if $\delta < (4 + 6z)/(7 + 10z)$. For the other values of δ, full cooperation in trade policy is feasible, that is, $\hat{\tau}^c > 0$. Furthermore, it is immediate to see that $\hat{\tau}^c < \tau^N$ for all values of δ and, therefore, some cooperation is achieved for any values of the parameters. Q. E. D.

Reference

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

Bagwell, K. and Staiger, R. (1990). A theory of managed trade. *American Economic Review*, *80*, 779–795.

Bagwell, K. and Staiger, R. (2001). Domestic policies, national sovereignty and international economic institutions. *The Quarterly Journal of Economics*, *116*, 519–562.

Copeland, B. (1990). Strategic interaction among nations: negotiable and non-negotiable trade barriers. *Canadian Journal of Economics*, *23*, 64–108.

Dixit, A. (1987). Strategic aspects of trade policy. In T.Bewley (Ed.), *Advances in Economic Theory: Fifth World Congress*. Cambridge University Press.

Ederington, J. (2001). International coordination of trade and domestic policies. *American Economic Review*, *91*, 1580–1593.

Ederington, J. (2002). Trade and domestic policy linkage in international agreements. *International Economic Review*, *43*, 1347–1367.

Green, E. and Porter, R. (1984). Noncooperative collusion under imperfect price information. *Econometrica*, *52*, 87–100.

Hungerford, T. (1991). Gatt : a cooperative equilibrium in a non-cooperative trading regime. *Journal of International Economics*, *31*, 357–369.

Limao, N. (2005). Trade policy, cross-border externalities and lobbies: do linked agreements enforce more cooperative outcomes?. *Journal of International Economics*, *67*, 175–199.

Markusen, J. (1975). International externalities and optimal tax structures. *Journal of International Economics*, *5*, 15–29.

Riezman, R. (1991). Dynamic tariffs with asymmetric information. *Journal of International Economics*, *30*, 267–283.

Spagnolo, G. (1996). Issue linkage, delegation and international policy cooperation. Tech. rep. 49.96, Fondazione Eni Enrico Mattei.