

Enforcement, Private Political Pressure and the GATT/WTO Escape Clause

by

Kyle Bagwell and Robert W. Staiger*

October 13, 2004

Revised: December 6, 2004

Abstract: We consider the design and implementation of international trade agreements when: (i) negotiations are undertaken and commitments made in the presence of uncertainty about future political pressures; (ii) governments possess private information about political pressures at the time that the agreement is actually implemented; and (iii) negotiated commitments can be implemented only if they are self-enforcing. We thus consider the design of self-enforcing trade agreements among governments that acquire private information over time. In this context, we provide equilibrium interpretations of GATT/WTO negotiations regarding upper bounds on applied tariffs and GATT/WTO escape clauses. We find that governments achieve greater welfare when they negotiate the optimal upper bound on tariffs rather than precise tariff levels; furthermore, when governments negotiate the optimal upper bound on tariffs, the observed applied tariffs often fall strictly below the bound. Our analysis also provides a novel interpretation of a feature of the WTO Safeguard Agreement, under which escape clause actions cannot be re-imposed in the same industry for a time period equal to the duration of the most recent escape clause action. We find that a dynamic usage constraint of this kind can raise the expected welfare of negotiating governments.

Bagwell: Columbia University (Kelvin J. Lancaster Professor of Economic Theory in the Department of Economics, and School of Business) and NBER. Staiger: University of Wisconsin at Madison (Department of Economics) and NBER. For helpful comments, we thank participants at The University of Chicago Conference on International Dispute Resolution. We thank the NSF (SES-0214021) for providing funding. All errors are ours.

I. Introduction

Efforts to implement international agreements typically take place in the presence of three broad features. First, negotiations are undertaken and commitments made in the presence of substantial *uncertainty* over the state of the world that will exist at the time the agreement is actually implemented. Second, at the time of implementation, governments are *privately informed* about specific features of the realized state of the world as these features apply to them. And third, negotiated commitments will be implemented successfully only if they are *self-enforcing*.

These three features are readily apparent in the context of international market access agreements negotiated within the GATT/WTO. For example, the market access commitments negotiated in the most recently completed (Uruguay) round of GATT/WTO negotiations were agreed to in 1994, but many of these commitments were not scheduled to be fully implemented until 10 years later: the elimination of the system of textile quotas known as the Multi-Fiber Agreement (MFA) is a prominent and instructive example. It seems fair to say that the state of the world that exists now that these commitments are due to be implemented was unknown at the time that governments originally signed on to these commitments. There also seem to be many examples (e.g., the growing political debate in the United States over the appropriate response to the cost inflicted on U.S. textile workers of the scheduled elimination of the MFA at the end of 2004) that confirm the observation that it is difficult for affected trading partners (e.g., China) to know with precision the extent of political pressure that another government actually faces at the time of implementation of its market access commitments. And finally, it seems widely acknowledged that a government can be expected to abide by commitments it negotiates within the GATT/WTO only if and for as long as it sees doing so to be in its self interest (i.e., GATT/WTO commitments are not meaningful unless they are self-enforcing).

In this paper, we suggest that these three broad features can help account for a number of broad characteristics of the design of the GATT/WTO and of the international commitments that are negotiated within the bargaining forum shaped by its rules. We emphasize two such characteristics that are central to the GATT/WTO. First, rather than negotiating precise tariff levels, the tariff

bindings with which governments anchor their negotiated market access commitments in the GATT/WTO define *upper bounds* above which a government agrees not to set its applied tariffs, and in practice the negotiated tariff bindings do not always appear to constrain the applied tariffs that governments select (i.e., governments sometimes set their applied tariffs at levels that are strictly below their associated GATT/WTO bindings). And second, rather than negotiating rigid tariff bindings to which governments agree to abide in all circumstances, the tariff bindings to which governments agree in GATT/WTO negotiations carry with them a variety of *escape clauses* that can be triggered under certain circumstances. As we demonstrate, these broad characteristics – which reflect choices made by the GATT/WTO member governments concerning both the design of their bargaining forum and the negotiating outcomes within this forum – follow naturally in an environment that exhibits the three features described above.

We undertake our analysis within a simple partial equilibrium two-country two-good trade model in which governments face political pressures from their import-competing producers and make choices over the levels of their import tariffs. Our formal model builds from Bagwell and Staiger (2001), but we assume here that the degree of political pressure faced by each government is a random variable that is determined (iid) each period according to a commonly known distribution function, and that each government privately observes the realization of its own political pressure in each period. We assume further for simplicity that the two governments can undertake negotiations only once, that they undertake these negotiations under a “veil of ignorance” (i.e., before observing present or future realizations of political pressure), and that they then make their tariff choices simultaneously in each subsequent period after privately observing their realized political pressure for the period and in light of the commitments they have negotiated.

We begin by considering negotiations over a rigid tariff binding that can be exogenously enforced. This tariff binding could either imply a *strong* commitment, or a *weak* commitment. Under a strong commitment, the negotiated tariff binding specifies the *precise* level at which a government must set its (applied) tariff. Under a weak commitment, the negotiated tariff binding specifies the *maximal* level at which a government may set its (applied) tariff. We characterize

negotiations over both, and establish that: (i) governments prefer negotiating commitments that take the form of weak bindings over strong bindings; (ii) when governments negotiate commitments that take the form of weak bindings, they choose to adopt tariff bindings which are higher than those they would choose if their negotiated commitments instead took the form of strong bindings; and (iii) when governments negotiate commitments that take the form of weak bindings, they choose to adopt tariff bindings which imply that governments with low realizations of political pressure will set their applied tariffs strictly below the bound level.

We next consider the enforcement of rigid tariff bindings, focusing on the case of weak bindings. We show that, when negotiated commitments are constrained to be self-enforcing: (i) the self-enforcement constraint does not bind provided that the rate at which governments discount the future is below a critical level; (ii) for discount rates just above this critical level, the negotiated tariff bindings must be higher – and the applied tariffs must therefore fall below the bound level with greater frequency – if the commitments are to remain self-enforcing; and (iii) there can be a critical discount rate above which governments are incapable of sustaining any commitments in a self-enforcing agreement, and are therefore destined to remain in the (one-shot) Nash tariff equilibrium.

Focusing on the high-discount-rate case, we then consider the value of introducing an escape clause in some form, as a way to enhance the ability of governments to benefit from self-enforcing tariff commitments and to expand the range of discount rates over which self-enforcing agreements are possible. We model an escape clause as allowing governments, in effect, to negotiate two bindings, one (lower) tariff to apply during “normal” times, and the other (higher) tariff to apply during “exceptional” times. Our first result here is negative: when bindings are weak, tariff commitments with an escape clause can offer no improvement over tariff commitments without an escape clause (i.e., the negotiation of a single rigid tariff). Intuitively, the problem is that each government’s private information over its realized political pressure will allow it to always claim sufficient pressure to exercise the higher escape clause binding, which it will surely do as it is then permitted to set its applied tariff at or below the escape clause binding, and so in effect there will only be one binding (the escape clause binding).

Our negative result is instructive: what is needed to make an escape clause potentially valuable to governments in this private-information setting is a cost associated with its use, so that exercising the escape clause might be made incentive compatible. When bindings are weak, the required cost is absent, and so an escape clause is worthless to governments. We proceed to consider several ways in which a cost associated with the use of an escape clause might be introduced, and ask whether an escape clause in this form could then be useful to governments.

A first approach is to insist that the escape clause binding is a strong commitment. The cost associated with the use of the escape clause is then the cost of being held to an applied tariff at the level of the escape clause binding. Evidently, this will be viewed as costly by a government in the event that the escape clause binding lies above that government's best-response tariff: this event in turn occurs for governments that experience relatively low realizations of political pressure. We show that this cost allows an incentive compatible escape clause to be designed, and that there exist escape clauses of this form that might enhance the value of government commitments. Of course, as we have observed, GATT/WTO commitments do not take this strong form. One interpretation of our result is then that introducing strong bindings into the GATT/WTO escape clause could enhance its value to governments. Appealing to the literature on trade agreements as devices to help governments make commitments to their own private sectors, we suggest as well that it may also be possible to offer an alternative interpretation of this result and of the cost associated with strong bindings that underlies it. Under this alternative interpretation, the escape clause binding is weak but a government operates in an environment of enhanced policy discretion whenever it invokes the escape clause. This policy discretion, in turn, can lead to tariff choices for the government that, while optimal ex-post, are sub-optimally protective ex-ante. Arguing in these broad terms, we suggest that the costs of government discretion may provide a way that incentive compatibility problems that would otherwise plague an escape clause with weak bindings might be mitigated.

A second approach to introduce a cost for invoking the escape clause is to maintain weak bindings but insist that a side-payment is made to the trading partner when a government uses the escape clause. We show that this too can allow an incentive compatible escape clause to be

designed, and that there exist escape clauses of this form that enhance the value of government commitments. Again we suggest two interpretations of our result. The first interpretation is literal: requiring that a government make cash payments to its trading partners when it invokes an escape clause can enhance the value of the escape clause to governments by addressing the incentive compatibility problems that would otherwise arise. As cash payments between governments have never been required within the GATT/WTO, this first interpretation indicates a possible direction for improvement of the design of the GATT/WTO escape clause. In this regard, we observe that the “quota rents” associated with Voluntary Export Restraints (VERs) are a form of cash payments made from the importing country to the exporting country, and so our result lends some support to GATT’s permissive stance regarding VERs and indicates some discord with the proscription of VERs introduced by the WTO. The second interpretation is more suggestive: we observe that it may be possible to interpret the compensation/retaliation provisions of the GATT/WTO escape clause as serving a similar purpose to the side-payments in our formal model, and thereby possible to interpret these provisions as critical for ensuring the incentive-compatible use of escape clause actions. We also discuss and interpret distinctions between the GATT and WTO-era escape clause rules from this perspective.

A third approach to introducing a cost to invoking the escape clause is to impose a dynamic use constraint on governments. We consider a particularly simple form: if a government uses the escape clause in this period, then it must wait a period before it can use the escape clause again. We demonstrate that this sort of constraint can work to address the incentive compatibility problem, and we show that there exist escape clauses of this form that enhance the value of government commitments. We suggest that this may be particularly relevant for the WTO safeguard provisions, which introduce a dynamic use constraint of this nature.

There are several strands of literature that relate to this paper. In the legal literature, Sykes (1991) is closest to our paper, and emphasizes many of the same themes that we emphasize here. In the economics literature, earlier papers by Bagwell and Staiger (1990, 2003), Feenstra and Lewis (1991), Rosendorff and Milner (2001) and Herzing (2004) are concerned specifically with the issue

of escape clauses in international agreements. Recent work by Martin and Vergote (2004) also explores some related themes in the context of anti-dumping policies. Also in the economics literature, the approach to analyzing collusion in the presence of private information developed in Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004) is methodologically related. We discuss the relationship of our results to these papers at later points in the paper.

The rest of the paper proceeds as follows. The next section presents the basic model and derives the incentive compatibility conditions in a static setting. Section III derives the optimal rigid tariff bindings when bindings are strong and when bindings are weak. Section IV introduces the repeated game to consider issues of enforcement. Section V considers the value of an escape clause in various forms. Section VI concludes. An Appendix supplies proofs not included elsewhere.

II. The Model

II.1 The Basic Setup

We begin by describing a simple two-country partial equilibrium world of trade in the presence of political considerations, analogous to that considered in Bagwell and Staiger (2001). In this world, the domestic (no *) country exports good y to the foreign (*) country in exchange for imports of good x from the foreign country.¹ With P_x and P_y denoting respectively the price of good x and good y in the domestic country market, domestic country demands for goods x and y are given by $D(P_x)=1-P_x$ and $D(P_y)=1-P_y$ respectively, while domestic country supplies of goods x and y are given by $Q_x(P_x)=P_x/2$ and $Q_y(P_y)=P_y$ respectively, with associated profit (producer surplus) functions $\pi_x(P_x)=(P_x)^2/4$ and $\pi_y(P_y)=(P_y)^2/2$. And with P_x^* and P_y^* denoting respectively the price of good x and good y in the foreign country market, the analogous expressions for foreign demands, supplies and profit (producer surplus) functions are $D^*(P_x^*)=1-P_x^*$, $D^*(P_y^*)=1-P_y^*$, $Q_x^*(P_x^*)=P_x^*$, $Q_y^*(P_y^*)=P_y^*/2$, $\pi_x^*(P_x^*)=(P_x^*)^2/2$ and $\pi_y^*(P_y^*)=(P_y^*)^2/4$. Domestic country imports of x and exports of y are then given by $M_x(P_x)\equiv D(P_x)-Q_x(P_x)=1-3P_x/2$ and $E_y(P_y)\equiv Q_y(P_y)-D_y(P_y)=2P_y-1$, respectively. Similarly, foreign country imports of y and exports of x are then given by

¹ As usual, the partial equilibrium model can be re-interpreted as a general equilibrium model in which there is an additional traded numeraire good in the background that consumers value in a quasi-linear fashion.

$M_y^*(P_y^*) \equiv D^*(P_y^*) - Q_y(P_y^*) = 1 - 3P_y^*/2$ and $E_x^*(P_x^*) \equiv Q_x^*(P_x^*) - D_x^*(P_x^*) = 2P_x^* - 1$, respectively.

The government of each country has a single policy instrument, namely, a specific tariff (τ for the domestic government and τ^* for the foreign government) that it can impose on its imports. Letting P_x^w be the price received by foreign exporters of x , we then have the following price expressions that must hold for any non-prohibitive τ : $P_x = P_x^w + \tau$ and $P_x^* = P_x^w$. Similarly, letting P_y^w be the price received by domestic country exporters of y , we then the following price expressions that must hold for any non-prohibitive τ^* : $P_y^* = P_y^w + \tau^*$ and $P_y = P_y^w$. We refer to P_x^w and P_y^w as the “world” (exporter) prices of x and y , respectively. Using these relationships, the market clearing world prices for x and y are determined by the respective market clearing conditions $M_x(P_x) = E_x^*(P_x^*)$ and $M_y^*(P_y^*) = E_y(P_y)$, yielding the market clearing prices $P_x^w(\tau) = [4 - 3\tau]/7$, $\hat{P}_x(\tau) = [4 + 4\tau]/7$, $P_y^w(\tau^*) = [4 - 3\tau^*]/7$ and $\hat{P}_y(\tau^*) = [4 + 4\tau^*]/7$ for $\tau < 1/6$ and $\tau^* < 1/6$. Trade in good x (good y) is prohibited when $\tau \geq 1/6$ ($\tau^* \geq 1/6$).

Finally, we define the government objectives. We follow Baldwin (1987) and assume that each government seeks to maximize a weighted sum of producer surplus, consumer surplus and tariff revenue in its economy, with a relatively greater weight on import-competing producer surplus reflecting political economy influences.² Specifically, let $\gamma \in [1, 7/4]$ denote the weight placed by the domestic government on its import-competing producer surplus. Then the welfare experienced by the domestic government on its import and export goods is defined as follows:

$$W_x \equiv \int_{\hat{P}_x}^1 D(P_x) dP_x + \gamma \pi_x(\hat{P}_x) + [\hat{P}_x - P_x^w] M_x(\hat{P}_x), \text{ and}$$

$$W_y \equiv \int_{\hat{P}_y}^1 D(P_y) dP_y + \pi_y(\hat{P}_y).$$

The domestic government welfare function is then the sum of these two terms. Similarly, let

²Grossman and Helpman (1994) provide micro-foundations to this interpretation.

$\gamma^* \in [1, 7/4]$ denote the weight placed by the foreign government on its import-competing producer surplus. Then the welfare experienced by the foreign government on its import and export goods is defined as:

$$W_y^* \equiv \int_{\hat{P}_y^*}^1 D^*(P_y^*) dP_y^* + \gamma^* \pi_y^*(\hat{P}_y^*) + [\hat{P}_y^* - P_y^w] M_y^*(\hat{P}_y^*), \text{ and}$$

$$W_x^* \equiv \int_{\hat{P}_x^*}^1 D^*(P_x^*) dP_x^* + \pi_x^*(\hat{P}_x^*).$$

The foreign government welfare function is then the sum of these two terms.

Notice that the domestic tariff choice only effects W_x and W_x^* , while the foreign tariff choice only effects W_y^* and W_y . In light of this, we may characterize Nash and efficient choices of τ with reference only to W_x and W_x^* , and similarly we may characterize Nash and efficient choices of τ^* with reference only to W_y^* and W_y . In addition, owing to the symmetry across goods x and y in our model, we may henceforth focus our characterizations on τ . Finally, our focus on political economy pressures in import-competing sectors only implies that only the domestic political economy parameter γ enters into the determination of Nash and efficient choices for τ . For future reference, we record the explicit expressions for W_x and W_x^* in terms of τ for given domestic political economy weight γ :

$$(1) \quad W_x(\tau; \gamma) = \frac{9+8\gamma}{98} + \frac{[8\gamma-5]}{49}\tau + \frac{2[2\gamma-17]}{49}\tau^2 \quad \text{and}$$

$$(2) \quad W_x^*(\tau) = \frac{25}{98} + \frac{9\tau^2}{49} - \frac{3\tau}{49}.$$

II.2 The Nash Tariff Choices and Efficient Tariff Choices

We next derive the Nash and efficient tariff choices. Using (1), the first-order necessary condition for the optimal unilateral choice of τ for the domestic government yields

$$(3) \quad \tau^N(\gamma) = \frac{8\gamma - 5}{4[17 - 2\gamma]}.$$

The second order condition is satisfied over the entire parameter range $\gamma \in [1, 7/4]$, and we note as well that τ^N is non-prohibitive over the parameter range $\gamma \in [1, 7/4]$, reaching the prohibitive level of $1/6$ when $\gamma = 7/4$. From (3), it may be confirmed that the domestic Nash tariff takes on a value of $1/20$ when $\gamma = 1$, reflecting the strictly positive optimal tariff of the national-income maximizing domestic government in this case, and that the domestic Nash tariff is increasing monotonically in γ , reflecting the desire for a higher tariff that is associated with increased weight placed on domestic import-competing producer surplus by the domestic government. Using (2), it may also be confirmed that W_x^* is strictly decreasing in $\tau = \tau^N(\gamma)$ over $\gamma \in [1, 7/4]$, indicating that the foreign government is strictly hurt by a further increase in the domestic tariff provided there is any export volume from the foreign country (i.e., provided $\tau < 1/6$).

The internationally efficient domestic tariff choice in our partial equilibrium setting is the domestic tariff that maximizes the sum of $W_x(\tau; \gamma) + W_x^*(\tau)$. Using (1) and (2), we have that

$$(4) \quad W_x(\tau; \gamma) + W_x^*(\tau) = \frac{9 + 8\gamma}{98} + \frac{[8\gamma - 5]}{49}\tau + \frac{2[2\gamma - 17]}{49}(\tau)^2 + \frac{25 + 18(\tau)^2 - 6\tau}{98}.$$

Using (4), the first-order necessary condition for the efficient choice of τ yields

$$(5) \quad \tau^E(\gamma) = \frac{4[\gamma - 1]}{25 - 4\gamma}.$$

Again the second order condition is satisfied over the entire parameter range $\gamma \in [1, 7/4]$. Comparing (3) and (5), we find that $\tau^N > \tau^E$ for $\gamma \in [1, 7/4)$, and we observe that $\tau^N = \tau^E$ for $\gamma = 7/4$. That is, the Nash tariff is inefficiently high as long as political economy forces do not drive the Nash tariff to a prohibitive level (in which case the Nash tariff is efficient). Figure 1 depicts the Nash and efficient domestic tariff levels over the entire parameter range $\gamma \in [1, 7/4]$. To focus our attention on political economy pressures that are consistent with strictly positive trade volumes, we henceforth restrict our

parameter range to $\gamma \in [\underline{\gamma}, \bar{\gamma}]$, where $\underline{\gamma} = 1$ and $\bar{\gamma} \in (\underline{\gamma}, 7/4)$.

II.3 Incentive Compatibility

We now assume that the political economy parameters γ and γ^* are each drawn independently from a common distribution with cumulative distribution function $H(\gamma)$ and density $h(\gamma) \equiv H'(\gamma)$ defined over the support $\gamma \in [\underline{\gamma}, \bar{\gamma}]$. We assume that each government privately observes its own political realization but not that of the other government.

In this private-information setting, an incentive compatibility problem may arise. For example, governments may wish to negotiate tariff bindings in such a way as to allow that a government's applied tariff may vary with its realized level of political pressure. Efficiency may be improved, for example, when a government applies a higher tariff when it experiences greater political pressure. The potential incentive compatibility problem is that a government might be tempted to apply a higher tariff, even when the political pressure that it actually experiences is low. Here we define and characterize the incentive constraints that rule out this kind of opportunistic behavior. For now, we continue to focus on a static model. In Section IV, we consider the enforcement of trade agreements and explain how the findings presented here extend to the repeated game context.

To this end, we define

$$G(\tau) \equiv \frac{9}{98} - \frac{5}{49}\tau - \frac{34}{49}\tau^2; \text{ and } f(\tau) \equiv \frac{4}{49} + \frac{8}{49}\tau + \frac{4}{49}\tau^2,$$

and observe that $f(\tau) > 0$ and $f'(\tau) > 0$ for $\tau \in [0, 1/6]$. Using (1), we may then write $W_x(\tau; \gamma)$ in the equivalent form $W_x(\tau; \gamma) = G(\tau) + \gamma f(\tau)$. Anticipating the possibility that an agreement might result in an applied tariff that is a function of political pressure, we let the applied tariff function be denoted by $\tau(\gamma)$. Given this tariff function, we may define the domestic government welfare (in good x) that is enjoyed when the realized political pressure is γ and the domestic government “announces” that

it faces $\hat{\gamma}$ (i.e., applies $\tau(\hat{\gamma})$):

$$(6) \quad \tilde{W}_x(\hat{\gamma}, \gamma) = G(\tau(\hat{\gamma})) + \gamma f(\tau(\hat{\gamma})).$$

The tariff function is incentive compatible if and only if, for all $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ and $\hat{\gamma} \in [\underline{\gamma}, \bar{\gamma}]$, the following is true:

$$(7) \quad \tilde{W}_x(\gamma, \gamma) \geq \tilde{W}_x(\hat{\gamma}, \gamma).$$

Incentive compatibility of the tariff function $\tau(\gamma)$ requires that the domestic government can't do better than announcing its realized political pressure truthfully. Using (6) this in turn requires that

$$\begin{aligned} \tilde{W}_x(\gamma, \gamma) &\geq \tilde{W}_x(\hat{\gamma}, \gamma) = \tilde{W}_x(\hat{\gamma}, \hat{\gamma}) + (\gamma - \hat{\gamma})f(\tau(\hat{\gamma})), \text{ and also} \\ \tilde{W}_x(\hat{\gamma}, \hat{\gamma}) &\geq \tilde{W}_x(\gamma, \hat{\gamma}) = \tilde{W}_x(\gamma, \gamma) + (\hat{\gamma} - \gamma)f(\tau(\gamma)). \end{aligned}$$

Rearranging yields

$$(8) \quad (\gamma - \hat{\gamma})f(\tau(\gamma)) \geq \tilde{W}_x(\gamma, \gamma) - \tilde{W}_x(\hat{\gamma}, \hat{\gamma}) \geq (\gamma - \hat{\gamma})f(\tau(\hat{\gamma})).$$

According to (8), a first requirement for incentive compatibility is that $f(\tau(\gamma))$ is non-decreasing in γ which, with $f'(\tau) > 0$ as noted above, implies in turn that $\tau(\gamma)$ must be non-decreasing in γ . Moreover, with $\tau(\gamma)$ monotonic, we then have that f is differentiable almost everywhere in γ , and therefore that $d\tilde{W}_x(\gamma, \gamma)/d\gamma$ is well defined.³ A second requirement for incentive compatibility is then that $d\tilde{W}_x(\gamma, \gamma)/d\gamma = f(\tau(\gamma))$: otherwise, by (6) the domestic government could benefit by misrepresenting its true political pressure. These two requirements for incentive compatibility are also sufficient (see, for example, Mas-Colell, Whinston and Green (1995)), and so we may state:

³For simplicity, we assume further that the applied tariff schedule $\tau(\gamma)$ is continuously differentiable, except perhaps at a finite number of points. We are thus able to include schedules with jumps.

Lemma 1 (IC): A tariff function $\tau(\gamma)$ is incentive compatible iff

- (i) $\tau(\gamma)$ is non-decreasing in γ , and
- (ii) $d\tilde{W}_x(\gamma, \gamma)/d\gamma = f(\tau(\gamma))$.

III. Rigid Tariffs

In this section we consider the possibility that governments might negotiate a single tariff binding that is used (in each sector) regardless of the realized political pressure that each government faces. By focusing on a rigid tariff binding, τ^R , we thus exclude for now the possibility that the governments might allow for safeguard exceptions, whereby a government is subjected to a higher binding under specific circumstances. We suppose that governments negotiate this tariff binding once and that the negotiation occurs before each observes its realized political pressure. As above, we continue to focus on a static model and defer discussion of the enforcement of trade agreements until Section IV.

We consider two forms that the tariff binding might take. Under a *strong binding*, the tariff binding defines the *precise tariff* that governments must apply under the agreement. Under a *weak binding*, the tariff binding defines the *maximum tariff* that governments can apply. An incentive compatibility problem clearly does not arise when bindings are strong, since the particular tariff that is applied does not depend upon the political pressure that a government experiences. Likewise, when bindings are weak, governments are sure to announce their types truthfully: only a government with sufficiently low political pressure could gain from applying a tariff below the binding, and such a government can apply its optimal tariff, $\tau^N(\gamma) < \tau^R$, by being honest. In each case, we look for the tariff binding that maximizes the ex-ante welfare of the two governments, and therefore the tariff binding that induces applied tariffs that maximize $E[W_x + W_x^*]$ where E denotes here the expectation operator and where expectations are taken over the domestic political pressure γ .⁴

⁴There is an equivalent choice of τ^{*R} which maximizes $E[W_y + W_y^*]$ with the expectations taken over the foreign political pressure γ^* . However, since the ex-ante symmetry of our setup ensures that the resulting choices of τ^R and τ^{*R} will be the same, we continue to characterize solutions in terms of the x -sector variables.

III.1 Strong Bindings

Under strong bindings, τ^R defines the precise tariff the governments must apply. To characterize the optimal τ^R in this case, we first define $\gamma^E(\tau^R)$ implicitly by $\tau^E(\gamma)=\tau^R$. In words, $\gamma^E(\tau^R)$ represents the marginal political realization, below which a government sets its tariff above the efficient level, and above which a government is constrained by the binding τ^R to set a tariff below the efficient level. Using (5), we calculate that

$$\gamma^E(\tau^R) = \frac{50\tau^R+8}{8[1+\tau^R]}.$$

Observe that $\gamma^{E'}(\tau^R)>0$. We next define the expected domestic political pressure in any period:

$$E\gamma \equiv \int_{\underline{\gamma}}^{\bar{\gamma}} \gamma dH(\gamma).$$

With τ^R independent of γ , we may now observe from (4) that $[W_x+W_x^*]$ is linear in γ . Hence, we have

$$(9) \quad E[W_x+W_x^*] = \frac{9+8E\gamma}{98} + \frac{[8E\gamma-5]}{49}\tau^R + \frac{2[2E\gamma-17]}{49}(\tau^R)^2 + \frac{25+18(\tau^R)^2-6\tau^R}{98},$$

which is identical to (4) with γ replaced by $E\gamma$.

The first order condition associated with (9) reduces to

$$(10) \quad E\gamma - \gamma^E(\tau^R) = 0.$$

In words, (10) indicates that the strong binding should be set at the level which is efficient for the expected realization of political pressure. Letting τ_s^R denote the rigid tariff with strong bindings that maximizes $E[W_x+W_x^*]$, we then have

$$(11) \quad \tau_s^R = \frac{4[E\gamma-1]}{25-4E\gamma},$$

which is by (5) simply $\tau^E(E\gamma)$.

We henceforth impose

(A1) $E\gamma > 5/4$.

Comparing (11) to (3), it may be checked that (A1) ensures that $\tau_s^R > \tau^N(\underline{\gamma})$. At later points in the paper we will appeal to the uniform distribution to characterize specific results. We observe here that (A1) is satisfied when γ is uniformly distributed over the support $[\underline{\gamma}, \bar{\gamma}]$ for $\bar{\gamma} \in [3/2, 7/4)$.

III.2 Weak Bindings

Under weak bindings, τ^R defines the maximum tariff that governments must apply. Thus, when τ^R is a weak binding, the applied tariff τ will depend on the realization of political pressure according to

$$\tau = \min\{\tau^N(\gamma), \tau^R\}.$$

To characterize the optimal τ^R in this case, we begin by defining $\gamma^N(\tau^R)$ implicitly by $\tau^N(\gamma) = \tau^R$. In words, $\gamma^N(\tau^R)$ represents the marginal political realization, below which a government is unconstrained by the binding τ^R and sets an applied tariff equal to $\tau^N(\gamma)$, and above which a government is constrained by the binding τ^R and applies its tariff at this level. Using (3), we calculate that

$$\gamma^N(\tau^R) = \frac{68\tau^R + 5}{8[1 + \tau^R]}.$$

Observe that $\gamma^N(\tau^R) > 0$, $\gamma^N(\tau^R) > 1 = \underline{\gamma}$ if and only if $\tau^R > \tau^N(\underline{\gamma}) = 1/20$, and $\gamma^N(\tau^R) < \bar{\gamma}$ if and only if $\tau^R < \tau^N(\bar{\gamma}) = [8\bar{\gamma} - 5]/[4(17 - 2\bar{\gamma})]$.

We are now prepared to define expected welfare as a function of a weak binding, τ^R . Our

definition reflects the fact that the applied tariff is $\tau^N(\gamma)$ when $\gamma \leq \gamma^N(\tau^R)$, while the applied tariff is equal to the binding, τ^R , when $\gamma \geq \gamma^N(\tau^R)$. We thus represent expected welfare under weak bindings as

$$(12) \quad E[W_x + W_x^* | \tau^R] = \int_{\underline{\gamma}}^{\gamma^N(\tau^R)} [W_x(\tau^N(\gamma); \gamma) + W_x^*(\tau^N(\gamma))] dH(\gamma) + \int_{\gamma^N(\tau^R)}^{\bar{\gamma}} [W_x(\tau^R; \gamma) + W_x^*(\tau^R)] dH(\gamma)$$

where we use notation that conditions on τ^R to denote that expected welfare is calculated for the case of a weak binding at τ^R .

Let us now define the expected political pressure conditional on being constrained by the binding τ^R . Denoting this conditional expectation by $E(\gamma | \tau^R)$, we have that

$$(13) \quad E(\gamma | \tau^R) = \frac{\int_{\gamma^N(\tau^R)}^{\bar{\gamma}} \gamma dH(\gamma)}{1 - H(\gamma^N(\tau^R))}.$$

Using (12) and (13), we may now write the effect of a small increase in τ^R on expected welfare when bindings are weak:

$$(14) \quad \frac{\partial E[W_x + W_x^* | \tau^R]}{\partial \tau^R} = \frac{[1 - H(\gamma^N(\tau^R))][8(1 + \tau^R)]}{49} \{E(\gamma | \tau^R) - \gamma^E(\tau^R)\}.$$

As (14) indicates, the impact of a small increase in τ^R when bindings are weak is only felt by the part of the distribution of political pressure to the right of (above) $\gamma^N(\tau^R)$. At an interior solution for the τ^R that maximizes $E[W_x + W_x^* | \tau^R]$, we must by (14) have

$$(15) \quad E(\gamma | \tau^R) - \gamma^E(\tau^R) = 0.$$

In words, (15) indicates that the weak binding should be set at the level which is efficient for the expected realization of political pressure conditional on being constrained by the binding τ^R .

We prove in the Appendix that an interior solution exists to the optimal weak binding problem. Denoting this solution by τ_w^R (which we assume for simplicity is unique), we may use the first order conditions (10) and (15) to conclude that $\tau_w^R > \tau_s^R$.⁵ That is, if governments negotiate a rigid tariff binding, they will set the binding higher if it is a weak binding than if it is a strong binding. In fact, it is direct to establish that $\tau^N(\bar{\gamma}) > \tau^E(\bar{\gamma}) > \tau_w^R > \tau_s^R > \tau^N(\underline{\gamma}) > \tau^E(\underline{\gamma}) = 0$. Figure 2 illustrates the position of τ_w^R and τ_s^R and these implied rankings. A second implication is that, when tariff bindings are weak, they will be set at a level which implies that a range of low-political-pressure realizations (i.e., for $\gamma \in [\underline{\gamma}, \gamma^N(\tau_w^R)]$) will cause governments to set their applied tariffs strictly below the bound level (see Figure 2). And a third implication is that governments prefer to negotiate weak bindings. To see this, note that τ_w^R gives higher expected welfare than would occur under weak bindings if τ_s^R were instead selected. But τ_s^R under weak bindings would deliver higher expected welfare than τ_s^R under strong bindings (the same welfare for $\gamma \geq \gamma^N(\tau_s^R)$, and higher welfare for $\gamma < \gamma^N(\tau_s^R)$ – see Figure 2). We summarize with

Proposition 1: Governments prefer negotiating commitments that take the form of weak bindings over strong bindings. When governments negotiate commitments that take the form of weak bindings, they choose to adopt tariff bindings which: (i) are higher than those they would choose if their negotiated commitments instead took the form of strong bindings; and (ii) imply that governments with low realizations of political pressure will set their applied tariffs strictly below the bound level.

It is instructive at this point to consider the case of the uniform distribution in more detail. Recall that (A1) is satisfied when γ is uniformly distributed over the support $[\underline{\gamma}, \bar{\gamma}]$ for $\bar{\gamma} \in [3/2, 7/4]$. We also have

$$E\gamma = \frac{1 + \bar{\gamma}}{2},$$

and

⁵This conclusion follows since (i) $\gamma^E(\tau^R)$ is strictly increasing, (ii) τ_w^R is interior and thus exceeds $\tau^N(\underline{\gamma})$, and (iii) $E(\gamma|\tau^R) > E\gamma$ for all $\tau^R > \tau^N(\underline{\gamma})$. Part (iii) is verified in the Appendix using (Z1) and (Z2).

$$E(\gamma|\tau^R) = \frac{\bar{\gamma} + \gamma^N(\tau^R)}{2}.$$

Using these expressions and (10), (15) and the expressions for $\gamma^E(\tau^R)$ and $\gamma^N(\tau^R)$, it follows that under the uniform distribution we have unique interior solutions for τ_s^R and τ_w^R given by

$$\tau_s^R = \frac{2[\bar{\gamma}-1]}{23-2\bar{\gamma}}; \text{ and } \tau_w^R = \frac{\bar{\gamma}-11/8}{4-\bar{\gamma}}.$$

It is now direct to confirm that the ranking $\tau^N(\bar{\gamma}) > \tau^E(\bar{\gamma}) > \tau_w^R > \tau_s^R > \tau^N(\underline{\gamma}) > \tau^E(\underline{\gamma}) = 0$ holds in the case of the uniform distribution, and the statements of Proposition 1 can be readily confirmed as well.

IV. Enforcement

Thus far, we have focused on a static model and ignored issues of enforcement. No external enforcement mechanism exists for the GATT/WTO; instead, trade agreements must be self-enforcing. In this section, we thus adopt a repeated game approach. We begin by describing the repeated game and the solution concept. Next, restricting our focus to weak bindings, we show that the applied tariffs induced by the optimal rigid tariff binding, τ_w^R , can be sustained in a self-enforcing agreement provided governments are sufficiently patient. Finally, we consider what can be done in the presence of governments that lack the required level of patience.

IV.1 The Repeated Game

The repeated game is informally defined as follows. In each period, each government privately observes its political economy parameter, where these draws are iid across governments and over time. The governments then simultaneously apply tariffs. The tariff selections are publically observed, and then the governments proceed to the next period. This process continues over an infinite number of periods, $t=1, \dots, \infty$.

As our solution concept, we use perfect public equilibria (Fudenberg, Levine and Maskin, 1994). We thus allow that a government's applied tariff strategy in any period t can depend upon

the government's current political economy parameter and the public history of past applied tariff selections. A government's period- t strategy does not depend upon its private information about its own past political economy parameters or the tariff strategies that it used in previous periods. Throughout the paper, we focus further on equilibria in which governments use symmetric strategies.⁶ Henceforth, we refer to symmetric perfect public equilibria simply as *equilibria*. In this section and most of the next section, we restrict attention to *stationary equilibria*. In a stationary equilibrium, along the equilibrium path, in every period a single function $\tau(\gamma)$ determines the tariff that is applied by a government with political economy parameter γ .

In our repeated game, an equilibrium must be robust against two types of deviations. A first kind of deviation is called an *on-schedule deviation*.⁷ Such a deviation occurs when a government observes that its current political economy parameter is γ and applies instead the equilibrium tariff that is called for when γ' is observed. Consider a stationary equilibrium. A government then undertakes an on-schedule deviation if it observes γ and yet applies the tariff $\tau(\gamma')$, where $\tau(\gamma')$ differs from $\tau(\gamma)$. Such a deviation is not observable, as a deviation, to the other government, which interprets the selection $\tau(\gamma')$ as indicating that γ' was observed. As a consequence, when a government observes that its current parameter is γ and considers an on-schedule deviation, the government simply compares the current-period welfare from selecting $\tau(\gamma)$ with that from selecting $\tau(\gamma')$. This is the same comparison that is defined in Section II.3 with respect to the incentive compatibility of an applied tariff schedule in the static game. Thus, for stationary equilibria of the repeated game, the applied tariff schedule $\tau(\gamma)$ does not invite an on-schedule deviation if and only if the schedule is incentive compatible in the static game.⁸

⁶Symmetric strategies do not depend upon governments' "names" and are thus interchangeable across governments. In other words, when strategies are symmetric, the current-period tariff strategies are permuted across governments when the governments' public histories of past tariff selections are permuted.

⁷For further discussion, see Athey, Bagwell and Sanchirico (2004).

⁸Of course, in non-stationary equilibria, on-schedule deviations must also be unattractive. In this case, though, the equilibrium may call for a different path of future play depending upon whether $\tau(\gamma)$ or $\tau(\gamma')$ is selected in the current period. As discussed in Section IV.4, the described relationship between on-schedule deviations in the repeated game and incentive compatibility in the static game is no longer valid.

A second kind of deviation is called an *off-schedule deviation*. A government undertakes an off-schedule deviation when it applies a tariff that is not called for in equilibrium, for any political economy parameter γ that the government might possibly have observed. In a stationary equilibrium, a government undertakes an off-schedule deviation if the government applies a tariff τ' where τ' differs from $\tau(\gamma)$ for all $\gamma \in [\underline{\gamma}, \bar{\gamma}]$. An off-schedule deviation is thus observable, as a deviation, to both governments. Once an off-schedule deviation is observed, play has moved off the equilibrium path, and the governments may then abandon cooperation and enter a punishment phase. For simplicity, we assume that following an off-schedule deviation governments revert to Nash play and select their Nash tariffs in all future periods.⁹ The off-schedule incentive constraint is thus analogous to a participation constraint.

Using these concepts, we may preview the analysis that follows. We assume that governments negotiate a bindings system at date zero, before learning their respective first-period political economy parameters. The bindings system then constrains applied tariffs, and we characterize the induced applied tariffs that correspond to equilibrium behavior. In the present section, we consider the enforcement of the tariffs that are applied under the optimal rigid tariff binding when bindings are weak. As noted in the previous section, the on-schedule incentive compatibility constraint for applied tariffs is trivially satisfied. The key issue is then whether governments are sufficiently patient to resist off-schedule deviations and thereby enforce the agreement. In the next section, we consider alternative binding systems in which the on-schedule constraint plays a more substantial role.

IV.2 Patient Governments

We characterize here the critical discount factor above which governments can maintain the applied tariffs induced by the optimal weak binding τ_w^R in a self-enforcing agreement. To this end, we define

⁹The reversion to Nash play is possible even in stationary equilibria, since stationarity is then required only for play along the equilibrium path. As noted below, other punishments would lead to qualitatively similar predictions.

$$EW^N \equiv E[W_x(\tau^N; \gamma) + W_x^*(\tau^N)],$$

and observe that $E[W_x(\tau_w^R; \gamma) + W_x^*(\tau_w^R) | \tau_w^R] > EW^N$.¹⁰ For simplicity, we assume that the alternative to cooperation is infinite reversion to Nash. In particular, if a government undertakes an off-schedule deviation by applying a tariff in excess of the weak binding, then a punishment phase is entered whereby Nash tariffs are applied for the remainder of time. Other punishments may lead to different values for the critical discount factor above which τ_w^R can be sustained, but the qualitative features of our results are robust to other punishments.

Let the per-period value of cooperation with any (not necessarily τ_w^R) rigid tariff under weak bindings be

$$\omega(\tau^R) \equiv E[W_x(\tau^R; \gamma) + W_x^*(\tau^R) | \tau^R] - EW^N,$$

where, as just observed, $\omega(\tau_w^R) > 0$. This per-period value of cooperation is received in each period by each government, since the countries and goods are symmetric and the levels of political pressure (types) are drawn iid. Let $\delta \in (0, 1)$ be the discount factor with which governments discount the future. That is, $\delta \equiv 1/[1+r]$ where r is the rate at which governments discount the future. Hence, a low δ is associated with impatient (high r) governments. The discounted future value of cooperation under τ^R is then

$$\frac{\delta}{1-\delta} \omega(\tau^R),$$

where $\delta/[1-\delta] \omega(\tau_w^R) > 0$. Figure 3 depicts all that can be said in general about $\delta/[1-\delta] \omega(\tau^R)$. In particular, as Figure 3 illustrates, $\delta/[1-\delta] \omega(\tau^R)$ takes its maximum value at τ_w^R and takes a value of zero at $\tau^N(\bar{\gamma})$. The slope of $\delta/[1-\delta] \omega(\tau^R)$ is also zero at τ_w^R and $\tau^N(\bar{\gamma})$.¹¹ But without further

¹⁰This follows because a weak binding at τ_w^R yields higher expected welfare than does a weak binding at $\tau^R = \tau^N(\bar{\gamma})$. The latter binding generates an expected welfare of EW^N .

¹¹A formal confirmation of the latter property is found in (Z6) in the Appendix.

restrictions on the distribution of γ , $\delta/[1-\delta]\omega(\tau^R)$ can be convex/concave, non-monotonic and even take on negative values between τ_w^R and $\tau^N(\bar{\gamma})$.

Next we consider the one-period incentive to cheat. When a government considers cheating on the agreement, it knows its realized political pressure (type) for that period. Fix any $\tau^R \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]$. If $\gamma \leq \gamma^N(\tau^R)$, then the government applies the tariff $\tau^N(\gamma)$, and so the incentive to cheat is zero. Next, suppose that $\gamma > \gamma^N(\tau^R)$. For these types, the binding lies below the Nash tariff, and so there exists an incentive to cheat by undertaking an off-schedule deviation and selecting $\tau^N(\gamma) > \tau^R$. We define this incentive as follows:

$$\Omega(\tau^R; \gamma) \equiv W_x(\tau^N(\gamma); \gamma) - W_x(\tau^R; \gamma) \text{ for } \gamma > \gamma^N(\tau^R).$$

Notice that $\Omega(\tau^R; \gamma) > 0$ and $\partial \Omega(\tau^R; \gamma) / \partial \gamma > 0$ over the domain $\gamma > \gamma^N(\tau^R)$. Therefore, the highest one-period incentive to cheat is associated with the highest realization of political pressure $\bar{\gamma}$, and so preventing this type from defecting from the agreement is necessary and sufficient for sustaining cooperation. The associated incentive constraint for a government with realized political pressure $\bar{\gamma}$ is

$$(16) \quad \Omega(\tau^R; \bar{\gamma}) \leq \frac{\delta}{1-\delta} \omega(\tau^R).$$

From the explicit expression for $\Omega(\tau^R; \bar{\gamma})$, given by

$$\Omega(\tau^R; \bar{\gamma}) = \frac{[13+8\bar{\gamma}]}{8[17-2\bar{\gamma}]} - \frac{9+8\bar{\gamma}}{98} - \frac{8\bar{\gamma}-5}{49} \tau^R + \frac{2[17-2\bar{\gamma}]}{49} (\tau^R)^2,$$

we may conclude that the one-period incentive to cheat for type $\bar{\gamma}$ exhibits the features depicted in Figure 4. In particular, $\Omega(\tau^R; \bar{\gamma})$ is decreasing and convex in τ^R for $\tau^R \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]$, and takes on a value of zero and has slope zero at $\tau^N(\bar{\gamma})$. In Figure 5, we plot $\Omega(\tau^R; \bar{\gamma})$ and $\delta/[1-\delta]\omega(\tau^R)$ together. As the figure illustrates, there exists a critical discount factor $\delta_w^R \in (0, 1)$ such that (16) is satisfied at $\tau^R = \tau_w^R$ for all $\delta \geq \delta_w^R$. Thus, the applied tariffs induced by the binding τ_w^R can be supported in a stationary

equilibrium, provided that $\delta \geq \delta_w^R$.

We now introduce some definitions so that our findings may be stated at a more formal level. First, for $\tau^R \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]$, the applied tariffs *induced* by a weak tariff binding τ^R are defined as the function $\tau(\gamma)$ such that $\tau(\gamma) = \tau^N(\gamma)$ for $\gamma \leq \gamma^N(\tau^R)$ and $\tau(\gamma) = \tau^R$ for $\gamma \geq \gamma^N(\tau^R)$. Second, for a given $\delta \in (0, 1)$, we say that a weak tariff binding τ^R is an *equilibrium binding* if a stationary equilibrium exists for the repeated game in which in each period governments select the applied tariffs induced by τ^R . Third, for a given $\delta \in (0, 1)$, we say that an equilibrium binding τ^R is an *optimal equilibrium binding* if no other equilibrium binding exists that yields higher expected welfare. We now have:

Proposition 2: There exists $\delta_w^R \in (0, 1)$ such that, for all $\delta \geq \delta_w^R$, τ_w^R is the optimal equilibrium binding.

IV.3 Impatient Governments

Let us now suppose that $\delta < \delta_w^R$. What can be done to try to preserve cooperation? A natural possibility is to allow the tariff binding τ^R to rise above τ_w^R . This will have two effects. First, as $\Omega(\tau^R|\bar{\gamma})$ is decreasing in τ^R for $\tau^R \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]$, raising τ^R will reduce $\Omega(\tau^R|\bar{\gamma})$, and this by itself helps to relax the incentive constraint (16). On the other hand, increasing τ^R above τ_w^R must eventually reduce $\delta/[1-\delta]\omega(\tau^R)$, since $\delta/[1-\delta]\omega(\tau^R)$ takes its maximum at τ_w^R and reaches a value of zero at $\tau^N(\bar{\gamma})$. Still, the level of $\delta/[1-\delta]\omega(\tau^R)$ is unaffected by small changes in τ^R around τ_w^R , and so for small changes the first effect must dominate. We may thus state:

Proposition 3: There exists $\underline{\delta} \in (0, \delta_w^R)$ such that for all $\delta \in [\underline{\delta}, \delta_w^R)$, the optimal equilibrium binding is characterized by $\tau^R(\delta)$ where $\tau^N(\bar{\gamma}) > \tau^R(\delta) > \tau_w^R$ and $\partial \tau^R / \partial \delta < 0$.

Proposition 3 captures the notion that less-patient governments may be forced to set a higher rigid tariff. One interesting implication in the presence of weak bindings is that, when cooperation is strained (due to impatience, or other analogous forces), our model predicts that tariff bindings will be higher, and applied tariffs will fall below the bound tariff level with greater frequency. It is also interesting to observe that the ability to maintain cooperation by raising the level of τ^R may not be

possible for δ sufficiently far below δ_w^R . The reason is that, as pointed out above, $\delta/[1-\delta]\omega(\tau^R)$ can be convex/concave, non-monotonic and even take on negative values between τ_w^R and $\tau^N(\bar{\gamma})$. Hence, the sustainable level of τ^R can fall smoothly to $\tau^N(\bar{\gamma})$ as δ falls toward zero, but it can also exhibit discontinuous jumps toward or all the way to $\tau^N(\bar{\gamma})$ along the way.

V. The Purpose and Design of the Escape Clause

It is now time to consider the value of introducing an escape clause in some form, as a way to enhance the ability of governments to benefit from self-enforcing tariff commitments and to expand the range of discount rates over which self-enforcing agreements are possible. We model an escape clause as allowing governments in effect to negotiate two bindings, one (lower) tariff to apply during “normal” times, and the other (higher) tariff to apply during “exceptional” times. Our first result, presented in the next subsection, is negative: when bindings are weak, tariff commitments with an escape clause can offer no improvement over tariff commitments without an escape clause (i.e., the negotiation of a single rigid tariff binding). Intuitively, the problem is that each government’s private information over its realized political pressure will allow it to always claim sufficient pressure to exercise the higher escape clause binding, which it will surely do as it is then permitted to set its applied tariff at or below the escape clause binding, and so in effect there will only be one binding (the escape clause binding).

However, our negative result is instructive: what is needed to make an escape clause potentially valuable to governments in this private-information setting is a cost associated with its use, so that exercising the escape clause might be made incentive compatible. When bindings are weak, the required cost is absent, and so an escape clause is worthless to governments. We proceed in subsequent subsections to consider several ways in which a cost associated with the use of an escape clause might be introduced, and ask whether an escape clause in this form could then be useful to governments.

V.1 Escape Clause with Weak Bindings

Throughout the paper, in any given period, we assume that governments apply their

respective tariffs after privately observing their respective types. In this context, the specification of a bindings system has the effect of selecting equilibrium behavior, since an applied tariff that violates a binding is regarded as an (off-schedule) deviation. In the present subsection, we consider a two-step binding system, under which an applied tariff in excess of the highest binding is regarded as an (off-schedule) deviation. Focusing on weak bindings, the question that we pose is whether such a system could ever induce applied tariffs that satisfy (on-schedule) incentive compatibility and yield higher expected welfare than can be achieved in a system with a single rigid tariff binding. As anticipated, our answer is negative, since the lower binding then has no real bite. The analysis is, however, instructive.

We establish a negative answer, making reference only to the on-schedule incentive constraints. It is thus sufficient to focus on the incentive compatibility constraints that arise in a static setting. Equivalently, we may imagine that the discount factor δ is sufficiently high, so that off-schedule constraints are assuredly satisfied (i.e., the highest binding is never exceeded).

We consider a two-step binding system. This system is described by a “breakpoint” type $\gamma^b \in (\underline{\gamma}, \bar{\gamma})$ such that types $\gamma \in [\underline{\gamma}, \gamma^b]$ bind at τ_1^R (i.e., apply a tariff no higher than τ_1^R), types $\gamma \in (\gamma^b, \bar{\gamma}]$ bind at τ_2^R (i.e., apply a tariff no higher than τ_2^R), and $\tau^N(\bar{\gamma}) \geq \tau_1^R, \tau_2^R$, where $\tau_2^R \neq \tau_1^R$.¹²

Recall from Lemma 1 that incentive compatibility can be achieved only if the *applied tariff* is (weakly) increasing in γ . We now consider what this implies for bindings. Suppose that $\tau_2^R < \tau_1^R$. Then from above we have $\tau^N(\bar{\gamma}) > \tau_2^R$. But then type $\bar{\gamma}$ would bind at τ_1^R , so that it could apply a higher tariff, closer to $\tau^N(\bar{\gamma})$. Thus, *bound tariffs* must also increase with γ : $\tau_2^R > \tau_1^R$. We may therefore state: There does not exist a system of two-step rigid tariffs with $\tau_2^R \leq \tau_1^R$.

There are now three remaining possibilities: (a) $\tau_1^R < \tau_2^R \leq \tau^N(\gamma^b)$; (b) $\tau^N(\gamma^b) \leq \tau_1^R < \tau_2^R$; and (c) $\tau_1^R < \tau^N(\gamma^b) < \tau_2^R$. We next observe that, given incentive compatibility, in any system of two-step rigid

¹²If one of the bindings exceeds $\tau^N(\bar{\gamma})$, then the two-step binding system clearly allows governments to apply their Nash tariffs, $\tau^N(\gamma)$, for all γ . This is the same outcome as would emerge under a single tariff binding at $\tau^N(\bar{\gamma})$.

tariffs, if $\tau_1^R \leq \tau^N(\gamma^b)$, then $\tau_2^R > \tau^N(\gamma^b)$. This is because otherwise, type $\gamma^b - \varepsilon$ would bind at τ_2^R , so that it could apply a higher tariff, closer to $\tau^N(\gamma^b - \varepsilon)$. This, though, rules out case (a) above. We therefore have: There does not exist a system of two-step rigid tariffs with $\tau_1^R < \tau_2^R \leq \tau^N(\gamma^b)$.

We next observe that case (b) above would leave no type constrained by τ_1^R , and therefore only τ_2^R would be constraining for types above a critical level of γ . But this would then be equivalent to a rigid tariff with weak bindings set at τ_2^R . We may therefore state: Any system of two-step rigid tariffs with $\tau_1^R \geq \tau^N(\gamma^b)$ is equivalent to a rigid tariff with weak bindings at $\tau^R = \tau_2^R$.

This leaves case (c) as the only possible avenue for improvement. But this case is also ruled out by incentive compatibility in light of the weak bindings: type $\gamma^b - \varepsilon$ would bind instead at τ_2^R , knowing that it could then almost apply its preferred tariff, since it could set $\tau = \tau^N(\gamma^b) < \tau_2^R$. Hence we have: There does not exist a system of two-step rigid tariffs with $\tau_1^R < \tau^N(\gamma^b) < \tau_2^R$.

Having exhausted all the possibilities, we conclude with:

Proposition 4: A system of two-step rigid tariffs cannot offer any improvement over a system with a single rigid tariff, when bindings are weak.

We close this subsection with two remarks. First, both private information and weak bindings are key for this result. If either were absent, then a two-step system could be designed which would improve upon the rigid tariff. And second, the problem is that there is no mechanism for preventing opportunistic use of the escape clause. This second observation has also been made by Rosendorff and Milner (2001) and Herzing (2004) in an environment that can be thought of as analogous to our weak-binding setting.¹³ In the next three subsections, we consider different ways of controlling opportunistic behavior by introducing a cost to the exercise of the escape clause so as to achieve

¹³Rosendorff and Milner (2001) and Herzing (2004) consider an environment in which bindings are strong but where an escape clause is interpreted as providing for the temporary absence of a binding. In the context of our discussion above, this interpretation carries with it the same implications as our weak-binding assumption.

incentive compatibility, and in this way we identify the possibility that an escape clause can become useful to governments.

V.2 Escape Clause with Strong Bindings

We next consider a hybrid system where the baseline bound tariff, τ_1^R , is subject to *weak* bindings, but the escape clause tariff binding, τ_2^R , can be used only if the binding is *strong* so that the government exercising the escape clause must actually *apply* τ_2^R . In this case, if $\tau_2^R > \tau_1^R$, a government that applies a tariff τ such that $\tau_2^R > \tau > \tau_1^R$ undertakes an off-schedule deviation and induces the Nash punishment in future periods. Intuitively, incentive compatibility might then be achieved, since a government now bears a cost when it exercises the escape clause (namely, being held to the precise tariff level τ_2^R) whenever τ_2^R lies strictly above the government's optimal unilateral tariff choice $\tau^N(\gamma)$.

To see how incentive compatibility could be achieved under this hybrid system, consider Figure 6. For purposes of illustration, we set $\tau_2^R = \tau^N(\bar{\gamma})$. With an appropriate choice of τ_1^R , type γ^b can be made the “breakpoint” type who is indifferent between τ_1^R and τ_2^R . Moreover, as observed in Section II.3, $\partial^2 W_x(\tau; \gamma) / \partial \tau \partial \gamma = f' > 0$, and so types higher than γ^b prefer τ_2^R while types lower than γ^b prefer τ_1^R . Of course, as Figure 6 illustrates, it is critical that the binding is strong; otherwise, the applied tariff (when τ_2^R is the binding) would be $\tau^N(\gamma^b)$, and no such τ_1^R could be found.

Notice that the achievement of an incentive compatible two-step arrangement entails a *cost*: when its type is high, a government sets its tariff *above* its best-response level (at least when $\gamma = \gamma^b$), hurting itself *and* its trading partner. Nevertheless, it is possible that governments could gain from such an arrangement. For example, suppose δ is below δ_w^R , so that governments are not patient enough to support the optimal rigid tariff with weak bindings, τ_w^R . Consider again the arrangement illustrated in Figure 6, only now set $\tau_1^R = \tau_w^R$, and let γ^b denote the implied breakpoint type (i.e., the type indifferent between $\tau_1^R = \tau_w^R$ and $\tau_2^R = \tau^N(\bar{\gamma})$). This arrangement leaves the incentive to cheat $\Omega(\tau^R; \gamma)$ unchanged for $\gamma \in [\underline{\gamma}, \gamma^b]$, but it *reduces* $\Omega(\tau^R; \gamma)$ for $\gamma \in (\gamma^b, \bar{\gamma}]$. Moreover, if $[1 - H(\gamma^b)]$ is sufficiently small, then the probability of realizations of political pressure at the level $\gamma \in (\gamma^b, \bar{\gamma}]$ in the

future will be small, and so the impact of this arrangement on $\omega(\cdot)$ will be approximately zero. In this case, the off-schedule incentive constraint is relaxed, and such an arrangement might succeed where a rigid adherence to τ_w^R would fail.¹⁴

As we have observed, the GATT/WTO escape clause does not take this strong form. Hence, one interpretation of our result is that it raises the possibility that introducing strong bindings into the GATT/WTO escape clause might enhance the value of the escape clause to governments. Appealing to the literature on trade agreements as devices to help governments make commitments to their own private sectors, we suggest as well that it may also be possible to offer an alternative interpretation of this result, and of the cost associated with strong bindings that underlies it. Under this alternative interpretation, the escape clause binding is *weak* but a government operates in an environment of enhanced policy *discretion* whenever it invokes the escape clause. This policy discretion, in turn, can lead to tariff choices for the government that, while optimal ex-post, are sub-optimally protective ex-ante.¹⁵ Under this alternative interpretation, then, a government that invokes the escape clause may find itself making tariff choices that lie strictly above its (ex-ante) optimal unilateral tariff choice, much as under the strong-binding arrangement illustrated in Figure 6. In these broad terms, we suggest that the costs of government discretion may provide a way that incentive compatibility problems that would otherwise plague an escape clause with weak bindings might be mitigated.

V.3 Escape Clause with Side Payments

We now consider a slight modification of our repeated game, in which governments may make transfers to each other at the time that they apply their respective tariffs. Maintaining our focus

¹⁴For a related argument in the context of a model of self-enforcing collusion, see Athey, Bagwell and Sanchirico (2004).

¹⁵For theoretical arguments that government discretion in trade policy can lead to sub-optimally high levels of protection, see, for example, Staiger and Tabellini (1987) and Maggi and Rodriguez-Clare (1998). For empirical evidence, see Staiger and Tabellini (1999), who compare the tariff binding choices made by the United States in the Tokyo Round of GATT negotiations with the U.S. tariff choices made under the GATT escape clause, and find that the added discretion associated with escape clause actions appears to lead to tariff decisions that are over-protective from an ex-ante perspective.

on stationary equilibria, we let $T(\gamma)$ be a transfer function that describes the transfer to be paid by the domestic government to the foreign government in exchange for the right to impose the tariff $\tau(\gamma)$. The transfer T has no direct efficiency implications, as it merely affects the allocation of surplus between the two governments, not the total amount of surplus.¹⁶ But the ability to make such transfers does have an important indirect effect on efficiency. Specifically, as we next demonstrate, if governments are sufficiently patient, then the fully efficient tariffs ($\tau^E(\gamma)$) can be implemented in a stationary equilibrium of the modified repeated game, where in each period along the equilibrium path the governments adopt a tariff-transfer scheme $\{\tau(\gamma), T(\gamma)\}$ with $\tau(\gamma) = \tau^E(\gamma)$.

To develop this result, we first show that an efficient tariff-transfer scheme can be constructed that satisfies the on-schedule incentive constraints (i.e., that is incentive compatible in a static game). Our second step is to suppose that governments use this scheme in each period and then show that no government can gain from cheating with an off-schedule deviation, if governments are sufficiently patient.

To begin, observe that the domestic government welfare (in good x) that is enjoyed when the realized political pressure is γ and the domestic government “announces” that it faces $\hat{\gamma}$ (i.e., applies $\tau(\hat{\gamma})$) will now be given by $\tilde{W}_x(\hat{\gamma}, \gamma) - T(\hat{\gamma})$. In analogy with (7) in Section II.3, the tariff-transfer scheme is incentive compatible if and only if, for all $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ and $\hat{\gamma} \in [\underline{\gamma}, \bar{\gamma}]$, the following is true:

$$(17) \quad \tilde{W}_x(\gamma, \gamma) - T(\gamma) \geq \tilde{W}_x(\hat{\gamma}, \gamma) - T(\hat{\gamma}).$$

In words, incentive compatibility of the tariff-transfer scheme $\{\tau(\gamma), T(\gamma)\}$ requires that the domestic government can't do better than announcing its realized political pressure truthfully. Under arguments exactly analogous to those in Section II.3 (see, e.g., Mas-Colell, Whinston and Green, 1995), the necessary and sufficient conditions for incentive compatibility of $\{\tau(\gamma), T(\gamma)\}$ are that (i) $\tau(\gamma)$ is non-decreasing in γ , and (ii) $d[\tilde{W}_x(\gamma, \gamma) - T(\gamma)]/d\gamma = f(\tau(\gamma))$. Using the Fundamental Theorem of

¹⁶As above, we focus on the x -sector, and symmetric tariff and transfer specifications are implied for the foreign government with respect to the y -sector.

Calculus to rewrite the second condition, we have:

Lemma 2 (IC): A tariff-transfer scheme $\{\tau(\gamma), T(\gamma)\}$ is incentive compatible iff

- (i) $\tau(\gamma)$ is non-decreasing in γ , and
- (ii) $\tilde{W}(\gamma, \gamma) - T(\gamma) = \tilde{W}(\underline{\gamma}, \underline{\gamma}) - T(\underline{\gamma}) + \int_{\underline{\gamma}}^{\gamma} f(\tau(\gamma)) d\gamma$.

Consider, then, $\tau^E(\gamma)$. Since $\tau^{E'}(\gamma) > 0$, condition (i) of Lemma 2 is satisfied. And we may define $T^E(\gamma; T(\underline{\gamma}))$ according to condition (ii) of Lemma 2, which using (6) leads to:

$$(18) \quad T^E(\gamma; T(\underline{\gamma})) \equiv [G(\tau^E(\gamma)) + \gamma f(\tau^E(\gamma))] - [G(\tau^E(\underline{\gamma})) + \underline{\gamma} f(\tau^E(\underline{\gamma}))] - \int_{\underline{\gamma}}^{\gamma} f(\tau^E(\gamma)) d\gamma + T(\underline{\gamma}).$$

We note by (18) that $\partial T^E(\gamma; T(\underline{\gamma})) / \partial \gamma > 0$. We may now state:

Lemma 3: For any $T(\underline{\gamma})$, the efficient tariff-transfer scheme $\{\tau^E(\gamma), T^E(\gamma; T(\underline{\gamma}))\}$ is incentive compatible.

We now turn to our second step and examine the off-schedule constraint. We argue that this consideration determines the intercept term $T(\underline{\gamma})$ in the efficient transfer function defined by (18). To establish this point, we note that, given symmetry, the per-period value of cooperation $\omega(\cdot)$ is independent of the transfer T . Hence, analogous to cooperation under the rigid tariff τ_w^R discussed in Section IV.1, the per-period value of cooperation under the efficient tariff transfer scheme, ω^E , is given by

$$\omega^E \equiv E[W_x(\tau^E(\gamma); \gamma) + W_x^*(\tau^E(\gamma))] - EW^N > 0.$$

Observe that $\delta/[1-\delta]\omega^E$ is then arbitrarily large as δ approaches 1.

Next we consider the one-period incentive to cheat. Again as in Section IV.1, when a government considers cheating on the agreement, it knows its realized political pressure (type) for

that period. A novel aspect now arises, however, in that a government can cheat on both its tariff selection and its transfer payment, *if* the called-for transfer payment is positive. To conform with the GATT/WTO practice of associating escape clause tariffs with the payment of compensation to affected exporters, we impose the restriction $T^E(\gamma) \geq 0$ for $\gamma \in [\underline{\gamma}, \bar{\gamma}]$, indicating that the tariff-imposing government pays a non-negative transfer to the exporting government. Under this restriction, cheating then means raising the import tariff to $\tau^N(\gamma)$ and making no transfer payment to the exporting government (with the tariff/transfer activity in the export sector unchanged in that period). Thus, the one-period incentive to cheat for type γ , $\Omega^E(\gamma)$, is given by

$$\Omega^E(\gamma; T(\underline{\gamma})) \equiv W_x(\tau^N(\gamma); \gamma) - [W_x(\tau^E(\gamma); \gamma) - T^E(\gamma; T(\underline{\gamma}))].$$

With these expressions, we may now state the off-schedule constraint for type γ :

$$(19) \quad \Omega^E(\gamma; T(\underline{\gamma})) \leq \frac{\delta}{1-\delta} \omega^E.$$

The lowest discount factor compatible with the off-schedule constraint for every type, denoted $\delta^E \in (0, 1)$, is found by selecting $T(\underline{\gamma})$ to minimize $\Omega^E(\gamma; T(\underline{\gamma}))$, subject to our constraint that $T^E(\gamma) \geq 0$ for $\gamma \in [\underline{\gamma}, \bar{\gamma}]$. Recall that, from (18), $\partial T^E(\gamma; T(\underline{\gamma})) / \partial \gamma > 0$. Therefore, since $\Omega^E(\gamma; T(\underline{\gamma}))$ is increasing in $T(\underline{\gamma})$, it follows that δ^E is found by setting $T(\underline{\gamma}) = 0$ and is defined implicitly by

$$\text{Max}_{\gamma \in [\underline{\gamma}, \bar{\gamma}]} \{W_x(\tau^N(\gamma); \gamma) - W_x(\tau^E(\gamma); \gamma) + T^E(\gamma; T(\underline{\gamma})=0)\} = \frac{\delta^E}{1-\delta^E} \omega^E.$$

We have thus established:

Proposition 5: There exists $\delta^E \in (0, 1)$ such that for all $\delta \in [\delta^E, 1]$, the fully efficient tariffs ($\tau^E(\gamma)$) can be implemented in a stationary equilibrium of the modified repeated game, where in each period along the equilibrium path the governments adopt the tariff-transfer scheme $\{\tau^E(\gamma), T^E(\gamma; T(\underline{\gamma})=0)\}$.

In effect, the transfer permits the efficient tariff function to be implemented for sufficiently patient

governments, with the intercept of the transfer $T(\gamma)$ being set to address the off-schedule incentive constraint and the slope of the transfer being set to address the on-schedule incentive constraint.

We suggest two interpretations of Proposition 5. The first interpretation is literal: requiring that a government make cash payments to its trading partners when it invokes an escape clause can enhance the value of the escape clause to governments by addressing the incentive compatibility problems that would otherwise arise, and in principle allow governments to implement fully efficient tariffs. As cash payments between governments have never been required within the GATT/WTO, this first interpretation indicates a possible direction for improvement of the design of the GATT/WTO escape clause. In this regard, we observe that the “quota rents” associated with Voluntary Export Restraints (VERs) are a form of cash payments made from the importing country to the exporting country, and so our result lends some support to GATT’s permissive stance regarding VERs and indicates some discord with the proscription of VERs introduced by the WTO Safeguards Agreement.

The second interpretation of Proposition 5 is more suggestive: we observe that it may be possible to interpret the compensation/retaliation provisions of the GATT/WTO escape clause as serving a similar purpose to the side-payments in our formal model, and thereby possible to interpret these provisions as critical for ensuring the incentive-compatible use of escape clause actions. Under this second interpretation, our result again indicates some discord -- or at least reason for caution -- with the WTO’s elimination of the compensation/retaliation provisions associated with escape clause actions (that meet certain criteria) for the first three years of their use. An interesting feature of the WTO Safeguard Agreement, however, is that in permitting governments to impose escape clause protection for up to three years without the need to compensate trading partners, a new constraint on the use of escape clause actions was also introduced: if a government imposes escape clause protection in an industry for a period of Y years, then it *cannot* reimpose escape clause protection in that industry for the next Y years. In the next subsection, we consider whether a *dynamic use constraint* of this nature can help address the incentive compatibility problem associated with the use of escape clause protection.

We close this subsection by briefly discussing the relationship between Proposition 5 and the findings of Feenstra and Lewis (1991), and also Bagwell and Staiger (1990, 2003), Rosendorff and Milner (2001), Herzing (2004) and Martin and Vergote (2004). Feenstra and Lewis report a finding similar to that reported in Proposition 5, in that the fully efficient tariff can be implemented with a tariff-transfer scheme analogous to ours. Our result here differs in two dimensions. First, Feenstra and Lewis impose an exogenous participation constraint, while our participation constraint is endogenously determined through the off-schedule constraint (19), i.e., Feenstra and Lewis do not address enforcement issues. And second, they restrict the transfer payment from importer to exporter to be no greater than the tariff revenue collected by the associated tariff, while we impose only that the transfer payment be non-negative. Proposition 5 is also somewhat related to the results of Bagwell and Staiger, with the most important distinction being that the results of Bagwell-Staiger are not derived in a private-information setting and hence are established without reference to an on-schedule incentive compatibility constraint. Rosendorff and Milner and also Herzing observe that requiring a fixed compensation payment for use of the escape clause (under which the binding is temporarily absent) can address the incentive compatibility problem, but these papers do not consider the optimal tariff-transfer scheme when the tariff and transfer can each depend on the announced state. Hence, the novelty of Proposition 5 relative to these papers is to characterize the optimal tariff-transfer scheme when both (off-schedule and on-schedule) constraints are operable. Finally, in recent work, Martin and Vergote also observe that transfers schemes can induce efficiency in repeated trade-policy games with private information. Their main focus, however, is on the use of retaliatory anti-dumping duties as a means of achieving greater cooperation among governments.

V.4 Escape Clause with Dynamic Use Constraint

As mentioned just above, the WTO Safeguard Agreement introduces a dynamic use constraint that applies to escape clause actions: once escape clause protection is removed, it cannot be re-imposed in the same industry for a period of time equal to the duration of the most recent escape clause action. Motivated by this constraint, we consider in this subsection a rule whereby the escape clause cannot be used two periods in a row. The effect of such a rule is to raise the opportunity cost of using the escape clause, and it might be conjectured that incentive compatibility

could then be achieved as a result. In this subsection, we confirm that this conjecture is indeed correct.

We consider a *two-state dynamic binding system* that involves two tariff bindings, $\underline{\tau}$ and $\bar{\tau}$, where $\tau^N(\bar{\gamma}) \geq \bar{\tau} > \underline{\tau} \geq \tau^N(\underline{\gamma})$. The first state occurs if, in the preceding period, a government applied a tariff τ such that $\tau \leq \underline{\tau}$. In this state, the government elected to bind at $\underline{\tau}$ in the preceding period. The second state arises if, in the preceding period, the government applied a tariff τ such that $\tau > \underline{\tau}$ and $\tau \leq \bar{\tau}$. Here, the government elected to bind at $\bar{\tau}$ in the previous period. In the first state, the government has flexibility in the current period: it can either bind again at $\underline{\tau}$ and thereby maintain the state, or it can bind at $\bar{\tau}$. In the second state, the government has no choice in the current period: having bound at $\bar{\tau}$ in the previous period, it must now bind at $\underline{\tau}$. To facilitate the comparison of our two-step dynamic binding system with the analysis above, we impose further that the bindings are positioned at or near the optimal rigid tariff binding when bindings are weak. Specifically, we consider $\underline{\tau} = \tau_w^R$ and $\bar{\tau} = \underline{\tau} + \varepsilon$ for $\varepsilon > 0$ and small. Finally, we assume that the governments begin the game in the first state.

As in Section V.1, this system is described by a “breakpoint” type. In the first state, where the government has flexibility, the breakpoint type is indifferent between binding at $\underline{\tau}$ (maintaining the state) and binding at $\bar{\tau}$ (changing the state). We denote the breakpoint type by $\gamma^* \in (\underline{\gamma}, \bar{\gamma})$. A first observation is that the system as defined requires that $\gamma^* > \gamma^N(\underline{\tau})$. If instead $\gamma^* \leq \gamma^N(\underline{\tau})$, then $\tau^N(\gamma^*) \leq \underline{\tau} < \bar{\tau}$. Type γ^* would then strictly prefer to bind at $\underline{\tau}$ by applying $\tau^N(\gamma^*)$, since it thereby enjoys its optimal tariff in the present and maintains future flexibility. A second observation is then that $\bar{\tau} < \tau^N(\gamma^*)$ for $\varepsilon > 0$ but sufficiently small.

With these observations in hand, we may refer to Figure 7. As illustrated there, in the two-state dynamic binding system, applied tariffs in the current period are *induced* as follows. In the first state when $\gamma \leq \gamma^*$, the government binds at $\underline{\tau}$; and it does so by applying its optimal tariff, $\tau^N(\gamma)$, if $\gamma \in [\underline{\gamma}, \gamma^N(\underline{\tau})]$, and by applying the tariff $\underline{\tau}$ if $\gamma \in [\gamma^N(\underline{\tau}), \gamma^*]$. In the first state when $\gamma \in (\gamma^*, \bar{\gamma}]$, the government binds at $\bar{\tau}$; and it does so by applying the tariff $\bar{\tau}$. Finally, in the second state, the

government must bind at $\underline{\tau}$; and it does so by applying its optimal tariff, $\tau^N(\gamma)$, if $\gamma \in [\underline{\gamma}, \gamma^N(\underline{\tau})]$, and by applying the tariff $\underline{\tau}$ if $\gamma \in [\gamma^N(\underline{\tau}), \bar{\gamma}]$.

We are now prepared to formally define and analyze the two-state dynamic binding system. There are three equations. The first two equations characterize the continuation value if last period's tariff binding was $\underline{\tau}$, which we denote by \underline{v} , and the continuation value if last period's tariff binding was $\bar{\tau}$, which we denote by \bar{v} . Defining $\underline{\tau}(\gamma) \equiv \min\{\tau^N(\gamma), \underline{\tau}\}$, it follows that \underline{v} and \bar{v} are defined by

$$(20) \quad \underline{v} = \int_{\underline{\gamma}}^{\gamma^*} W_x(\underline{\tau}(\gamma); \gamma) dH(\gamma) + H(\gamma^*) \delta \underline{v} + \int_{\gamma^*}^{\bar{\gamma}} W_x(\bar{\tau}; \gamma) dH(\gamma) + (1 - H(\gamma^*)) \delta \bar{v}$$

and

$$(21) \quad \bar{v} = \int_{\underline{\gamma}}^{\bar{\gamma}} W_x(\underline{\tau}(\gamma); \gamma) dH(\gamma) + \delta \bar{v}.$$

The third equation characterizes the required indifference of type γ^* :

$$(22) \quad W_x(\underline{\tau}; \gamma^*) + \delta \underline{v} = W_x(\bar{\tau}; \gamma^*) + \delta \bar{v}$$

This three-equation system has three unknowns: \underline{v} , \bar{v} and γ^* . Our next task is to use these equations to prove that $\gamma^* \in (\gamma^N(\underline{\tau}), \bar{\gamma})$ does exist, and that it is unique.

Using $W_x(\tau; \gamma) = G(\tau) + \gamma f(\tau(\gamma))$, we may manipulate (20)-(22) to obtain

$$(23) \quad \gamma^* - (E^* - \gamma^*) \delta (1 - H(\gamma^*)) = -[\bar{G} - \underline{G}] / [\bar{f} - \underline{f}],$$

where $\bar{G} \equiv G(\bar{\tau})$, $\underline{G} \equiv G(\underline{\tau})$, $\bar{f} \equiv f(\bar{\tau})$, $\underline{f} \equiv f(\underline{\tau})$, and

$$E^* \equiv \frac{\int_{\underline{\gamma}}^{\bar{\gamma}} \gamma dH(\gamma)}{1 - H(\gamma^*)}.$$

We prove in the Appendix the following:

Lemma 4: There exists a unique $\gamma^* \in (\gamma^N(\underline{\tau}), \bar{\gamma})$ solving (20)-(22).

We now consider the existence of an equilibrium in which, along the equilibrium path, the governments select the applied tariffs induced by the two-state dynamic binding system. We note that such an equilibrium is non-stationary: along the equilibrium path, current play is determined by the state and thus the history of play.¹⁷ As (22) reveals, the on-schedule incentive constraint, which occurs only in the first state, now reflects a dynamic comparison and is no longer equivalent to the incentive compatibility constraint of the static game. With $f'(\tau) > 0$ and $\bar{\tau} > \underline{\tau}$, we see that indifference of type γ^* ensures that higher (lower) types prefer to bind at $\bar{\tau}$ ($\underline{\tau}$). Thus, the on-schedule constraint is satisfied by construction. As before, we specify that an off-schedule deviation results in a reversion to Nash tariffs in all future periods. An off-schedule deviation occurs if a government applies a tariff in excess of $\bar{\tau}$; likewise, a government undertakes an off-schedule deviation if it is in the second state and yet applies a tariff in excess of $\underline{\tau}$. For $\varepsilon > 0$ and sufficiently small, the applied tariffs induced by the two-state dynamic binding system generate an expected discounted welfare that is strictly higher than that under Nash tariffs.¹⁸ Thus, if $\delta < 1$ is sufficiently large, the future cost of an off-schedule deviation is then sure to overwhelm any immediate gain from cheating. We thus have:

Proposition 6: There exists $\varepsilon^* > 0$ and $\delta^* < 1$ such that, for all $\varepsilon \in (0, \varepsilon^*)$ and $\delta \in (\delta^*, 1)$, there exists a non-stationary equilibrium for the repeated game in which, along the equilibrium path, governments

¹⁷The associated perfect public equilibrium is symmetric, since governments adopt the same strategy for the game. The equilibrium is not strongly symmetric, however: if the governments are in different states at a given point in time, they will enjoy different continuation values. See Athey and Bagwell (2001) for a related analysis in which different continuation values are used to provide truth-telling incentives in a repeated game with private information.

¹⁸When $\varepsilon > 0$ is small, the applied tariffs induced by the two-state dynamic binding system generate approximately the same expected discounted welfare as do the applied tariffs that are induced by the weak tariff binding, τ_w^R . As we show in Proposition 2, the latter give higher expected discounted welfare than would the applied tariffs induced by a weak tariff binding at $\tau^R = \tau^N(\bar{\gamma})$. But a weak tariff binding at $\tau^R = \tau^N(\bar{\gamma})$ induces the application of Nash tariffs and thus generates the expected discounted welfare of Nash play.

select the applied tariffs induced by the two-state dynamic binding system in which $\underline{\tau}=\tau_w^R$ and $\bar{\tau}=\underline{\tau}+\varepsilon$.

According to Proposition 6, sufficiently patient governments can support the applied tariffs induced by the two-state dynamic binding system as a non-stationary equilibrium of the repeated game. A remaining question is whether this two-state arrangement generates expected discounted game welfare that exceeds that which arises under the optimal weak tariff binding at τ_w^R . Our finding is as follows:

Proposition 7: Suppose that the distribution function, $H(\gamma)$, is uniform. Suppose further that the discount factor $\delta < 1$ is sufficiently large that equilibria for the repeated game exist in which governments select the applied tariffs induced by (i) the optimal weak tariff binding τ_w^R , and (ii) the two-state dynamic binding system in which $\underline{\tau}=\tau_w^R$ and $\bar{\tau}=\underline{\tau}+\varepsilon$. If $\varepsilon > 0$ is sufficiently small, then the two-state dynamic binding system generates strictly higher expected discounted game welfare than does the optimal weak tariff binding, τ_w^R .

Thus, under further distributional assumptions, if governments are sufficiently patient, the escape clause captured by the dynamic two-state binding system offers a strict gain for government welfare.¹⁹ In this way, the results of this sub-section provide an economic foundation for the dynamic usage constraint found in the WTO Safeguard Agreement. Proposition 7 is proved in the Appendix. Intuitively, when the distribution function is sufficiently close to uniform, γ^* is not too far below $\gamma^E(\tau_w^R)$, so that the move from $\underline{\tau}=\tau_w^R$ to $\bar{\tau}$ is a move *toward* $\tau^E(\gamma)$ for most of the probability distribution above γ^* (see Figure 7).

VI. Conclusions

We consider the design and implementation of international trade agreements, when (i) negotiations are undertaken and commitments made in the presence of uncertainty about future political pressures; (ii) governments possess private information about political pressures at the time

¹⁹The result holds also if the distribution function is sufficiently close to uniform. Proposition 7 holds, for example, under the power distribution, whereby $H(\gamma)=[(\gamma-\underline{\gamma})/(\bar{\gamma}-\underline{\gamma})]^\alpha$, if $\alpha > 0$ is sufficiently close to unity.

that the agreement is actually implemented; and (iii) negotiated commitments can be implemented successfully only if they are self-enforcing. Our analysis thus offers insights with respect to the optimal design of self-enforcing trade agreements when governments acquire private information over time.

We also provide equilibrium interpretations of specific features of GATT/WTO negotiations. In particular, we offer interpretations of GATT/WTO negotiations regarding upper bounds on applied tariffs and GATT/WTO escape clauses. In our model, governments achieve greater welfare when they negotiate upper bounds on tariffs rather than precise tariff levels; furthermore, when governments negotiate the optimal upper bound on tariffs, the observed applied tariffs often fall strictly below the bound. This is consistent with the observation of governments that often set their applied tariffs strictly below their associated GATT/WTO bindings. Our analysis also provides a novel interpretation of a feature of the WTO Safeguard Agreement, under which escape clause actions cannot be re-imposed in the same industry for a time period equal to the duration of the most recent escape clause action. We find that a dynamic use constraint of this kind can raise the expected welfare of negotiating governments.

Our formal analysis does not capture the potential role that a dispute settlement body might play in this context. At several points in the paper, we suggest that GATT's approach to escape clause actions seems more in line with the broad message of our formal analysis than does the approach taken by the WTO. However, one possible interpretation is that with the creation of a more effective dispute settlement body, the WTO seeks to rely more on its newly created legal system and investigatory powers to force the direct revelation of conditions that would justify escape clause actions than was possible under GATT's weaker legal environment. Whether this is indeed practical, and if so whether the WTO environment has achieved an appropriate mix between the use of legal investigative measures and the kind of incentives against escape clause abuse that we have highlighted and studied in this paper is an interesting and important question that we leave for future research.

References

- Athey, Susan and Kyle Bagwell, "Optimal Collusion with Private Information," *The Rand Journal of Economics*, vol. 32, no. 3, Autumn 2001, pp. 428-465.
- Athey, Susan, Kyle Bagwell and Chris Sanchirico, "Collusion with Price Rigidity," *The Review of Economic Studies*, vol. 71, no. 2, 2004, pp. 317-349.
- Bagwell, Kyle and Robert W. Staiger, "A Theory of Managed Trade," *American Economic Review*, vol. 80, September 1990, pp. 779-795 .
- Bagwell, Kyle and Robert W. Staiger, "Reciprocity, Non-Discrimination and Preferential Agreements in the Multilateral Trading System," *European Journal of Political Economy* vol. 17, no. 2, June 2001, pp. 281-325.
- Bagwell, Kyle and Robert W. Staiger, "Protection over the Business Cycle," *Advances in Economic Analysis & Policy*, Vol. 3, No. 1, 2003, Article 3, 1-43.
- Baldwin, Richard, "Politically Realistic Objective Functions and Trade Policy," *Economics Letters*, vol. 24, pp. 287-90.
- Feenstra, Robert C. and Tracy R. Lewis, "Negotiated Trade Restrictions with Private Political Pressure," *Quarterly Journal of Economics*, vol. 196, no. 4, November 1991, pp. 1287-1307.
- Grossman, Gene M. and Elhanan Helpman, "Protection for Sale," *American Economic Review*, vol. 84, September 1994, pp. 833-50.
- Herzing, Mathias, "The Escape Clause as a Contingent Contract," October 2004, mimeo, Institute for International Economic Studies.
- Maggi, Giovanni and Andres Rodriguez-Clare, "The Value of Trade Agreements in the Presence of Political Pressure," *Journal of Political Economy*, vol. 106, June 1998, pp. 574-601.
- Martin, Alberto and Wouter Vergote, "A Case for Strategic Antidumping," 2004, Columbia University.
- Mas-Colell, A., M. D. Whinston and J. R. Green, **Microeconomic Theory**, Oxford University Press, 1995.
- Rosendorff, Peter and Helen Milner, "The Optimal Design of International Trade Institutions: Uncertainty and Escape," *International Organization*, vol. 55, no. 4, Autumn 2001, pp. 829-857.

Staiger, Robert W. and Guido Tabellini, "Discretionary Trade Policy and Excessive Protection," *American Economic Review*, vol. 77, March 1987, 1265-77.

Staiger, Robert W. and Guido Tabellini, "Do GATT Rules Help Governments Make Domestic Commitments?," *Economics and Politics*, vol. 11, 1999, pp. 109-44.

Sykes, Alan O., "Protectionism as a 'Safeguard': A Positive Analysis of the GATT 'Escape Clause' with Normative Speculations," *University of Chicago Law Review* 58, 1991, pp. 255-305.

Appendix

In this Appendix, we first prove that an interior solution exists to the optimal weak binding problem leading up to Proposition 1.

Proof (interior solution): To establish the existence of an interior solution for the optimal weak binding problem, we begin by recording some properties of the conditional expectation, $E(\gamma | \tau^R)$, defined in (13):

$$E(\gamma | \tau^N(\underline{\gamma})) = E\gamma; \quad (\text{Z1})$$

$$\frac{\partial E(\gamma | \tau^R)}{\partial \tau^R} > 0 \text{ for all } \tau^R \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]; \quad (\text{Z2})$$

$$\lim_{\tau^R \rightarrow \tau^N(\bar{\gamma})} E(\gamma | \tau^R) = \bar{\gamma}; \text{ and} \quad (\text{Z3})$$

$$\lim_{\tau^R \rightarrow \tau^N(\bar{\gamma})} \frac{\partial E(\gamma | \tau^R)}{\partial \tau^R} = \frac{1}{2} \frac{\partial \gamma^N(\tau^N(\bar{\gamma}))}{\partial \tau^R}. \quad (\text{Z4})$$

The proof of (Z1) is immediate, and (Z2) also follows from straightforward calculations. A single application of L'Hopital's Rule yields (Z3), and (Z4) can be shown to follow from repeated applications of L'Hopital's Rule.

Next, we consider the derivative given in (14). First, at $\tau^R = \tau^N(\underline{\gamma})$, we have from (Z1) that $E(\gamma | \tau^R) = E\gamma$. Further, we find that $\gamma^E(\tau^R) = 5/4$ and $\gamma^N(\tau^R) = \underline{\gamma}$. Thus, by (A1), we have

$$\frac{\partial E[W_x + W_x^* | \tau^R]}{\partial \tau^R} > 0 \text{ at } \tau^R = \tau^N(\underline{\gamma}). \quad (\text{Z5})$$

Second, at $\tau^R = \tau^N(\bar{\gamma})$, we have from (Z3) that $E(\gamma | \tau^R)$ is finite, and of course $\gamma^E(\tau^R)$ is also finite. Since $\gamma^N(\tau^R) = \bar{\gamma}$, $H(\gamma^N(\tau^R)) = 1$ and thus

$$\frac{\partial E[W_x + W_x^* | \tau^R]}{\partial \tau^R} = 0 \text{ at } \tau^R = \tau^N(\bar{\gamma}). \quad (\text{Z6})$$

Third, at $\tau^R = \tau^N(\bar{\gamma})$, we may differentiate (14) and use (Z3), $H(\gamma^N(\tau^R)) = 1$ and (Z4), where the latter ensures that $\frac{\partial E(\gamma | \tau^R)}{\partial \tau^R}$ is finite, to find

$$\frac{\partial^2 E[W_x + W_x^* | \tau^R]}{\partial (\tau^R)^2} = H'(\bar{\gamma}) \frac{\partial \gamma^N(\tau^N(\bar{\gamma}))}{\partial \tau^R} \frac{2[25 - 4\bar{\gamma}]}{49} [\tau^N(\bar{\gamma}) - \tau^E(\bar{\gamma})].$$

Under our assumption that $\bar{\gamma} < 7/4$, it follows that $\tau^N(\bar{\gamma}) > \tau^E(\bar{\gamma})$, and so

$$\frac{\partial^2 E[W_x + W_x^* | \tau^R]}{\partial (\tau^R)^2} > 0 \text{ at } \tau^R = \tau^N(\bar{\gamma}). \quad (\text{Z7})$$

By (Z6) and (Z7), we may conclude that $E[W_x + W_x^* | \tau^R]$ obtains a local minimum at $\tau^R = \tau^N(\bar{\gamma})$. Given (Z5), it follows that $E[W_x + W_x^* | \tau^R]$ has an interior maximizer, $\tau_w^R \in (\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma}))$. Finally, at the interior maximizer, (15) must hold. Given (Z2) and (Z3), it follows that $E(\gamma | \tau_w^R) < \bar{\gamma}$. At $\tau^R = \tau^E(\bar{\gamma})$, $\gamma^E(\tau^R) = \bar{\gamma}$. Since $\gamma^E(\tau^R)$ is strictly increasing, the interior maximizer, τ_w^R , must satisfy $\tau_w^R \in (\tau^N(\underline{\gamma}), \tau^E(\bar{\gamma}))$.

Proof of Lemma 4: We may re-write (4) as

$$0 = \delta\bar{\gamma} + \beta(\bar{\tau} + \underline{\tau}) - \gamma^*(1 + \delta) - \delta \int_{\gamma^*}^{\bar{\gamma}} H(\gamma) d\gamma \equiv RHS(\gamma^*), \quad (Z8)$$

where

$$\beta(\bar{\tau} + \underline{\tau}) \equiv \frac{-[\bar{G} - \underline{G}]}{[\bar{f} - \underline{f}]} = \frac{5 + 34(\bar{\tau} + \underline{\tau})}{8 + 4(\bar{\tau} + \underline{\tau})}. \quad (Z9)$$

Using (Z9), we observe that $\beta'(\bar{\tau} + \underline{\tau}) > 0$ and $\beta(2\underline{\tau}) = \gamma^N(\underline{\tau})$. For $\bar{\tau} > \underline{\tau}$ with $\tau^N(\bar{\gamma}) \geq \bar{\tau}$ and $\underline{\tau} \geq \tau^N(\underline{\gamma})$, it follows that $\beta(\bar{\tau} + \underline{\tau}) > \beta(2\underline{\tau}) \geq \beta(2\tau^N(\underline{\gamma})) = \gamma^N(\tau^N(\underline{\gamma})) = \underline{\gamma}$, and $\beta(\bar{\tau} + \underline{\tau}) < \beta(2\bar{\tau}) \leq \beta(2\tau^N(\bar{\gamma})) = \gamma^N(\tau^N(\bar{\gamma})) = \bar{\gamma}$.

We may easily verify that $RHS'(\gamma^*) < 0$. Further, simple calculations confirm that

$$\begin{aligned} RHS(\bar{\gamma}) &= \beta(\bar{\tau} + \underline{\tau}) - \bar{\gamma} < 0, \text{ and} \\ RHS(\underline{\gamma}) &= \delta[E\underline{\gamma} - \underline{\gamma}] + \beta(\bar{\tau} + \underline{\tau}) - \underline{\gamma} > 0. \end{aligned}$$

This establishes the existence of a unique solution $\gamma^* \in (\underline{\gamma}, \bar{\gamma})$ to (Z8).

We next show that this solution satisfies $\gamma^* > \gamma^N(\underline{\tau})$. To this end, we compute that

$$RHS(\gamma^N(\underline{\tau})) = [\beta(\bar{\tau} + \underline{\tau}) - \gamma^N(\underline{\tau})] + \delta[1 - H(\gamma^N(\underline{\tau}))][E(\gamma | \underline{\tau}) - \gamma^N(\underline{\tau})].$$

Note that $\delta[1 - H(\gamma^N(\underline{\tau}))] > 0$, since $\delta > 0$ and $\underline{\tau} < \bar{\tau} \leq \tau^N(\bar{\gamma})$. Next, $\beta(\bar{\tau} + \underline{\tau}) > \beta(2\underline{\tau}) = \gamma^N(\underline{\tau})$. Finally, let $y(\underline{\tau}) \equiv E(\gamma | \underline{\tau}) - \gamma^N(\underline{\tau})$. Observe that $y(\tau^N(\underline{\gamma})) = E\underline{\gamma} - \underline{\gamma} > 0$ by (Z1), and $y(\tau^N(\bar{\gamma})) = E(\gamma | \tau^N(\bar{\gamma})) - \bar{\gamma} = 0$ by (Z3). Further, simple calculations reveal that

$$y'(\underline{\tau}) = \frac{\partial \gamma^N(\underline{\tau})}{\partial \underline{\tau}} \left\{ \frac{H'(\underline{\tau})}{1 - H(\underline{\tau})} y(\underline{\tau}) - 1 \right\}.$$

Thus, if $y(\underline{\tau}) \leq 0$ for $\underline{\tau} < \tau^N(\bar{\gamma})$, then $y(\tau^N(\bar{\gamma})) < 0$, which is a contradiction. Hence, $y(\underline{\tau}) > 0$ for all $\underline{\tau} \in [\tau^N(\underline{\gamma}), \tau^N(\bar{\gamma})]$. We may conclude that $RHS(\gamma^N(\underline{\tau})) > 0$. The solution thus satisfies $\gamma^* > \gamma^N(\underline{\tau})$.

Proof of Proposition 7: Referring to (15), we may express the first order condition as follows:

$$\frac{\partial E[W_x + W_x^* | \tau^R]}{\partial \tau^R} = \int_{\gamma^N(\tau^R)}^{\bar{\gamma}} \frac{\partial}{\partial \tau^R} [W_x(\tau^R; \gamma) + W_x^*(\tau^R)] dH(\gamma) = 0 \quad (\text{Z10})$$

at $\tau^R = \tau_w^R$. Let

$$N(\tau; \gamma) \equiv \frac{\partial}{\partial \tau^R} [W_x(\tau^R; \gamma) + W_x^*(\tau^R)] H'(\gamma). \quad (\text{Z11})$$

Using (Z11), we may re-write (Z10) as

$$\int_{\gamma^N(\tau^R)}^{\bar{\gamma}} N(\tau^R; \gamma) d\gamma = 0 \quad (\text{Z12})$$

at $\tau^R = \tau_w^R$.

When the distribution function is uniform, $H'(\gamma) = 1/[\bar{\gamma} - \underline{\gamma}]$. Thus, with $W_x(\tau^R; \gamma) = G(\tau^R) + \gamma f(\tau^R)$, we have

$$\frac{\partial N(\tau^R; \gamma)}{\partial \gamma} = f'(\tau^R)/[\bar{\gamma} - \underline{\gamma}] > 0. \quad (\text{Z13})$$

Using (Z12) and (Z13), we thus conclude: for all $\tilde{\gamma} \in (\gamma^N(\tau_w^R), \bar{\gamma})$, there exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon \in [0, \bar{\varepsilon}]$,

$$\int_{\tilde{\gamma}}^{\bar{\gamma}} N(\tau_w^R + \varepsilon; \gamma) d\gamma > 0. \quad (\text{Z14})$$

We now compare the efficiency gain offered by an alternative applied tariff schedule relative to the applied tariff schedule induced by the weak tariff binding at τ_w^R . Fixing $\tilde{\gamma} \in (\gamma^N(\tau_w^R), \bar{\gamma})$ and $\bar{\varepsilon}$ that satisfy (Z14), we define the alternative schedule by $\tau^A(\gamma) = \tau^N(\gamma)$ for $\gamma \in [\underline{\gamma}, \gamma^N(\tau_w^R)]$, $\tau^A(\gamma) = \tau_w^R$ for $\gamma \in [\gamma^N(\tau_w^R), \tilde{\gamma}]$, and $\tau^A(\gamma) = \tau_w^R + \bar{\varepsilon}$ for $\gamma \in [\tilde{\gamma}, \bar{\gamma}]$. The schedules differ only for $\gamma > \tilde{\gamma}$. Thus, by

(Z14), the efficiency gain is positive:

$$\begin{aligned}
& \int_{\tilde{\gamma}}^{\bar{\gamma}} \{ [W_x(\tau_w^R + \bar{\varepsilon}; \gamma) + W_x^*(\tau_w^R + \bar{\varepsilon})] - [W_x(\tau_w^R; \gamma) + W_x^*(\tau_w^R)] \} dH(\gamma) \\
= & \int_{\tilde{\gamma}}^{\bar{\gamma}} \int_{\tau_w^R}^{\tau_w^R + \bar{\varepsilon}} \frac{\partial}{\partial \tau^R} (W_x(\tau^R; \gamma) + W_x^*(\tau^R)) H'(\gamma) d\tau^R d\gamma \\
= & \int_{\tau_w^R}^{\tau_w^R + \bar{\varepsilon}} \int_{\tilde{\gamma}}^{\bar{\gamma}} N(\tau^R; \gamma) d\gamma d\tau^R \\
= & \int_0^{\bar{\varepsilon}} \int_{\tilde{\gamma}}^{\bar{\gamma}} N(\tau_w^R + \varepsilon; \gamma) d\gamma d\varepsilon > 0.
\end{aligned}$$

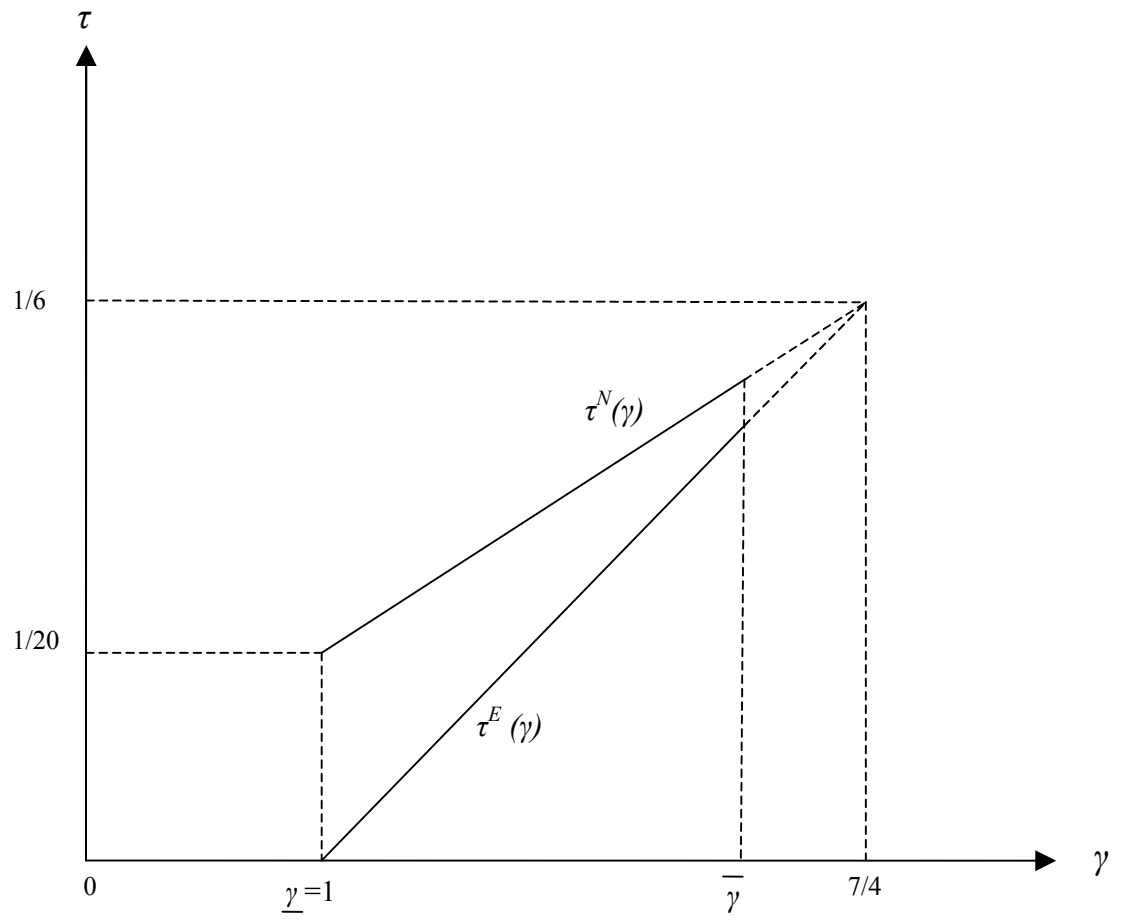


Figure 1

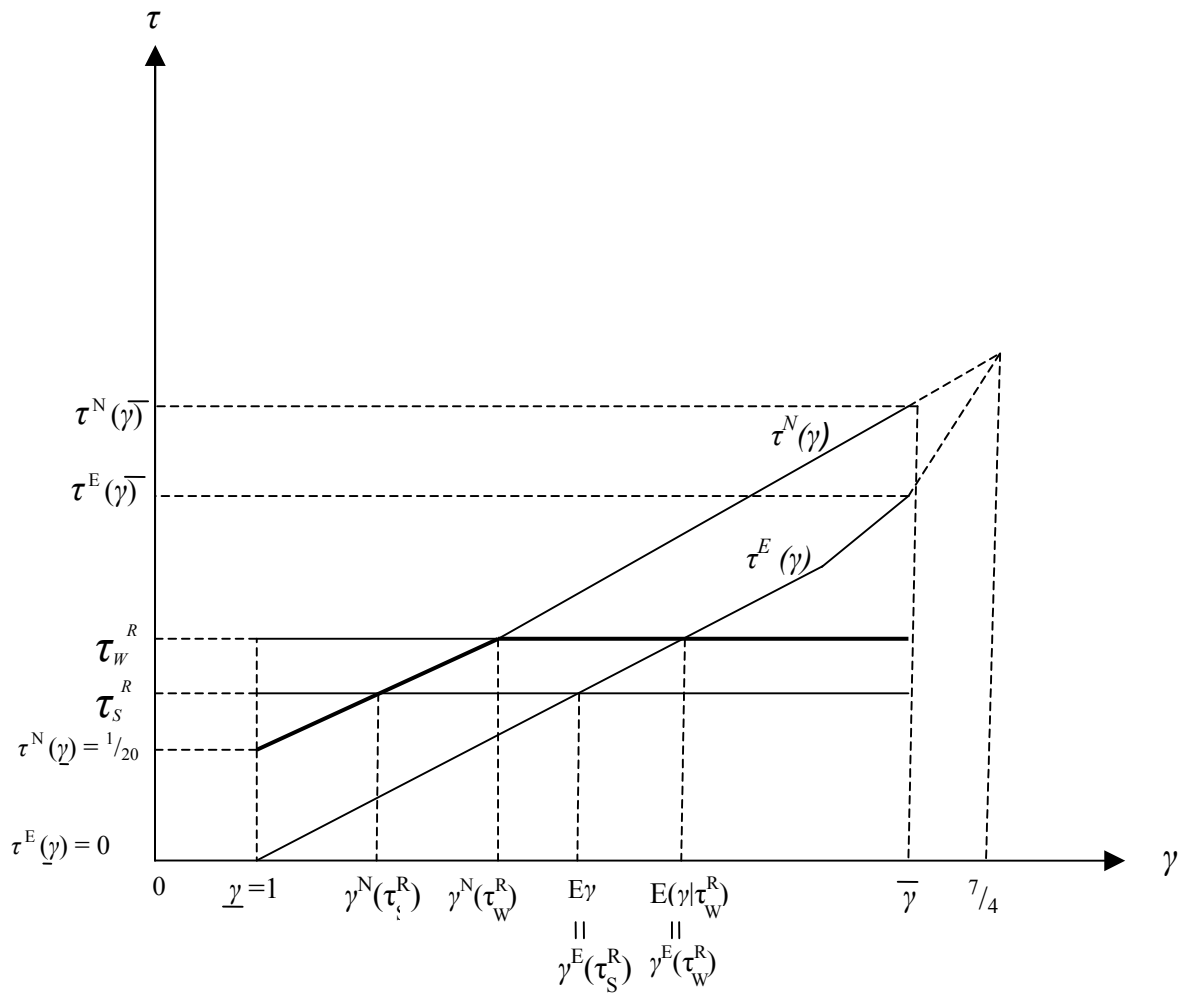


Figure 2

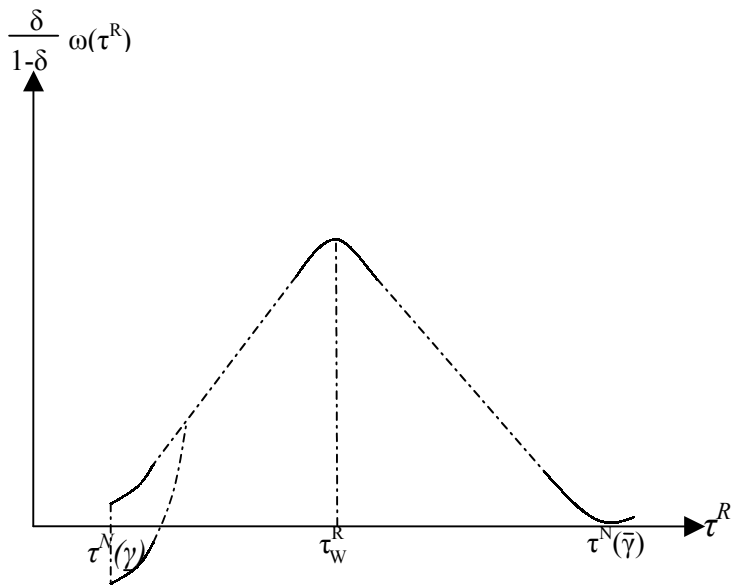


Figure 3

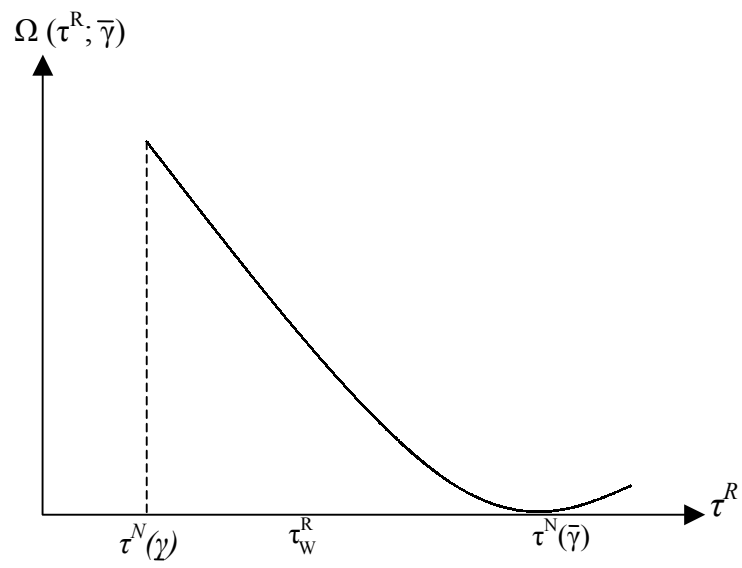


Figure 4

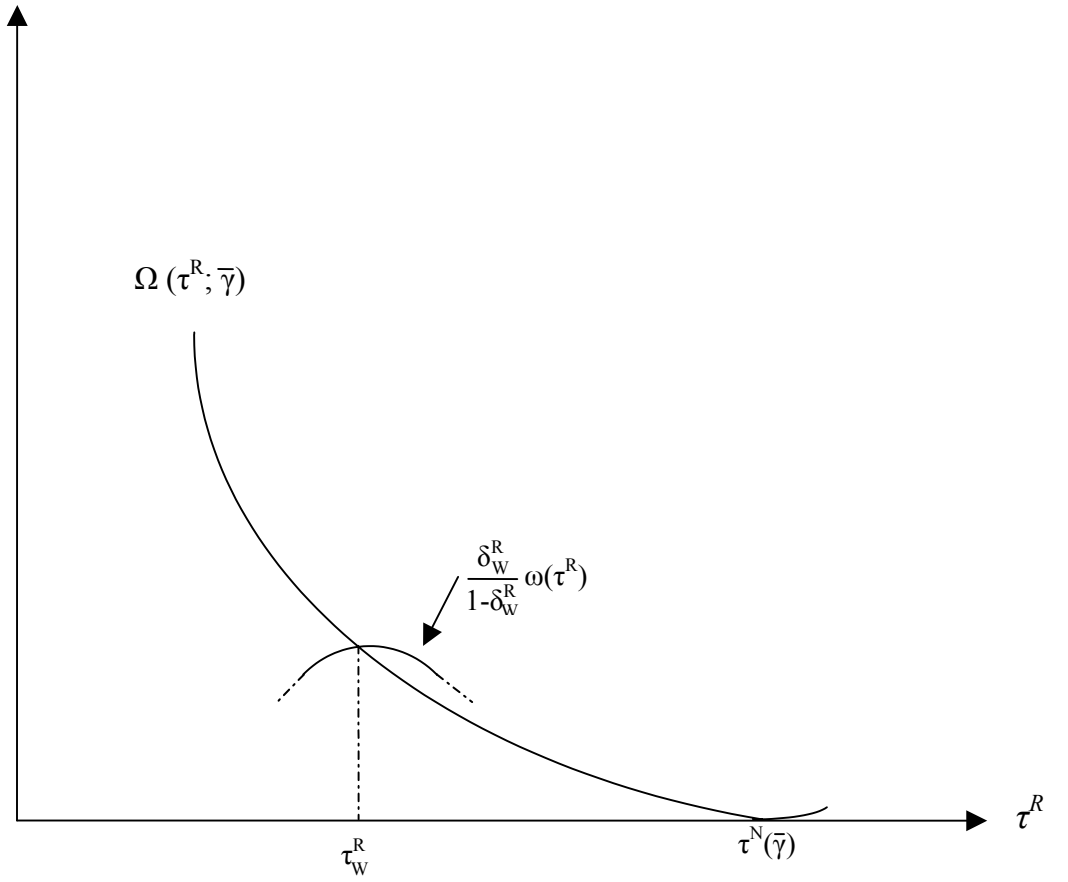


Figure 5

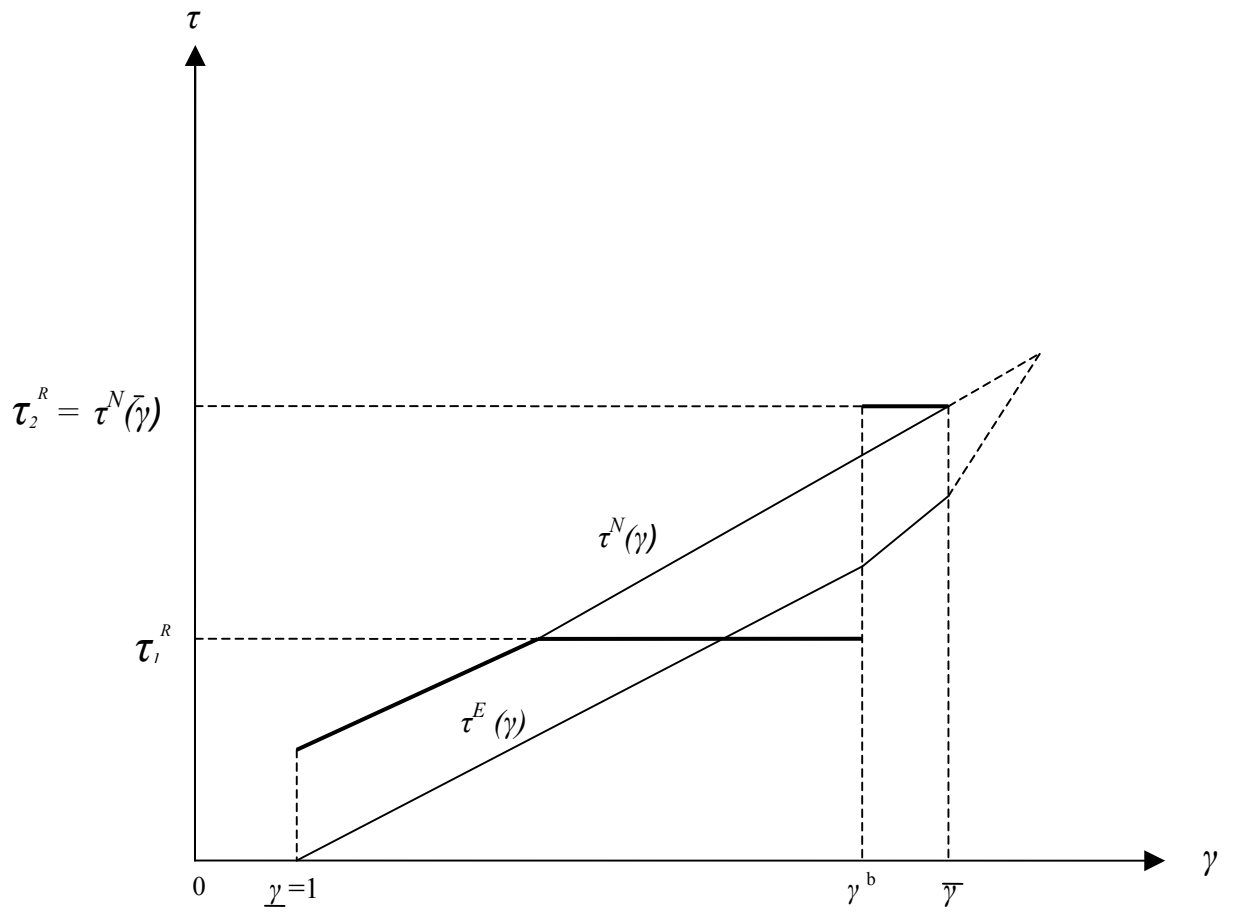


Figure 6

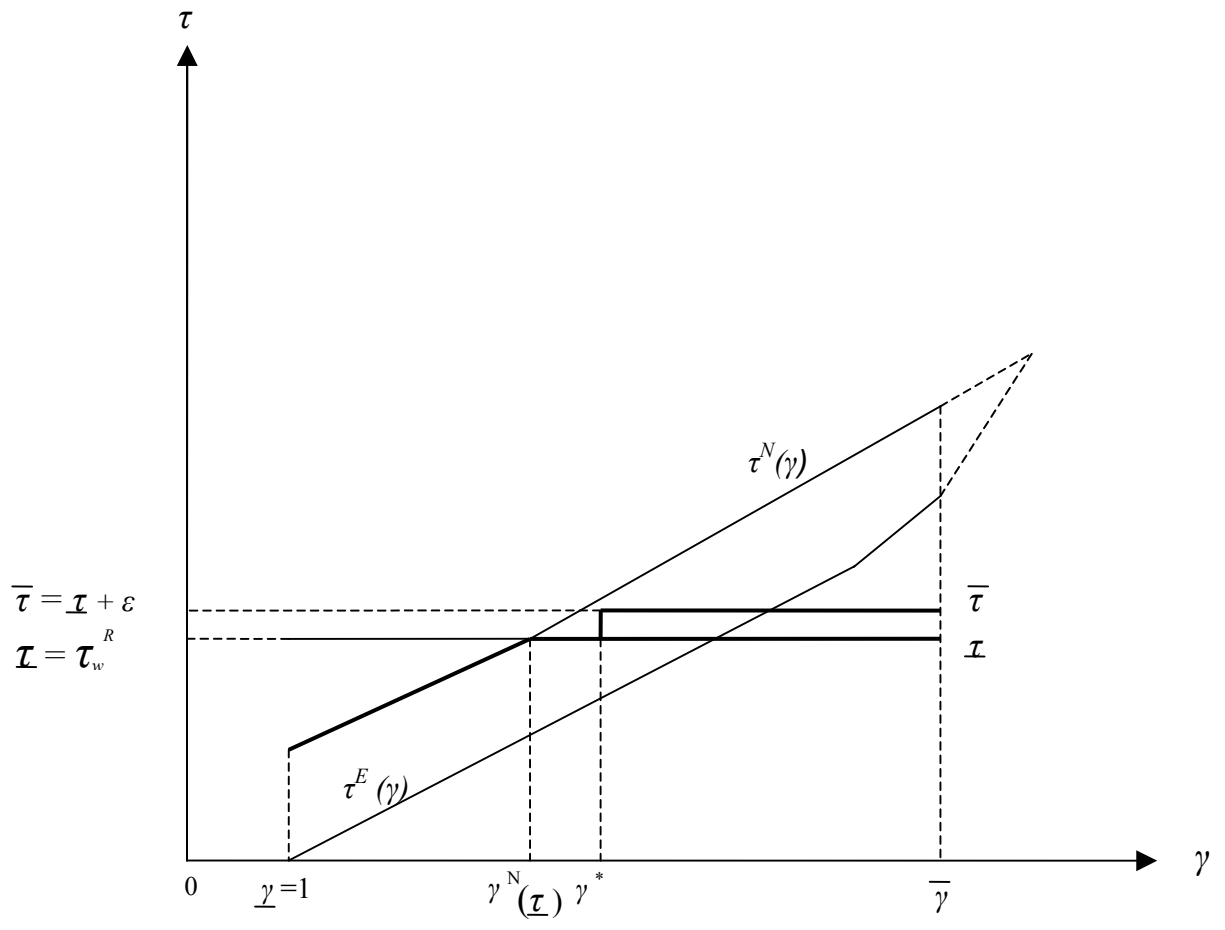


Figure 7