Self-Enforcing Trade Agreements and Private Information

Kyle Bagwell

Preliminary Draft: October 15, 2007

Abstract

This paper considers self-enforcing trade agreements among privately informed governments. A trade agreement that uses weak bindings (i.e., maximal tariff levels) is shown to offer advantages relative to a trade agreement that uses strong bindings (i.e., precise tariff levels). Consistent with practice, the theory also predicts that governments sometimes apply tariffs that are strictly below their bound rates. When private information is persistent through time, an enforcement "ratchet effect" is identified: a government reveals that it is "weak," and thus that it is unlikely to retaliate in an effective manner, when it applies a low tariff. This effect suggests that a government with a low type may "pool" at an above-optimal tariff, in order to conceal weakness. It also suggests a new information-based theory of gradualism in trade agreements.

Affiliation: Kelvin J. Lancaster Professor of Economic Theory (Department of Economics) and Professor of Finance and Economics (School of Business), Columbia University; and NBER. The helpful comments of Yeon-Koo Che, Robert Staiger and seminar participants at UBC and the financial support of the NSF are gratefully acknowledged.

1. Introduction

When a government imposes an import tariff, the tariff lowers the world price at which foreign exporters sell and thus induces a terms-of-trade loss for the foreign country. A trade agreement that features reciprocal liberalization can then be mutually beneficial for governments by raising trade volumes to more efficient levels, where efficiency is measured relative to the political-economic preferences of governments.[1] A cooperative trade agreement of this kind, however, can be achieved only if it is self-enforcing.[2] In other words, cooperation can be achieved only if each government perceives that its short-term benefit of cheating is outweighed by the expected discounted future cost of the consequent deterioration in cooperation.

When governments design a self-enforcing trade agreement, a further complication is that they are uncertain about the political pressures that they will face in the future. One implication is that governments cannot know with certainty the tariff levels that will be efficient from their joint perspective in future periods. Of course, if political pressures were publicly observable, then governments could design a state-contingent agreement, in which they agree to apply the tariffs that are efficient given whatever future preferences they may hold. Broad movements in political pressures may be publicly observable; however, at any given point in time, each government is likely to possess some private information about the extent of its political pressures. Private information introduces an incentive-compatibility problem. If the trade agreement is designed to allow a government to apply a high tariff when it reports a high degree of political pressure from its import-competing industry, then even a government with low pressure may be tempted to "lie" and apply the high tariff, as it thereby enjoys a terms-of-trade gain. The presence of private information may also interact with the requirement that the trade agreement be self-enforcing. This interaction arises when, as seems plausible, the pressures that a government privately observes have persistent components. In that case, if a government applies a low tariff, then another government may perceive that the former government faces persistent and low political pressures. The former government then may be perceived as "weak" and thus unlikely to retaliate aggressively should another government cheat. Once this inference is made, the trade agreement may fail to be self-enforcing.

In sum, when governments design a trade agreement while recognizing that they will each possess private information in the future, they may seek to lower average tariffs, so as to address the terms-of-trade externality, and to allow each government some discretion when applying its tariff, so as to facilitate better matching between a government's applied tariff and its (privately observed) political pressures. Finally, governments must recognize as well that, if the benefits of better matching are realized, then a government's applied tariff may reveal its

[1] For further discussion of the terms-of-trade rationale for trade agreements between governments with political-economic preferences, see Bagwell and Staiger (1999, 2001, 2002) and Grossman and Helpman (1995).

[2] The view that trade agreements must be self-enforcing is advanced by Bagwell and Staiger (1990, 2002), Dixit (1987) and Maggi (1999), for example.

private information. If private information is persistent in nature, this revelation could in turn interact with the self-enforcement constraint, by altering another government's perception as to the severity of retaliation that it would face in the future were it to cheat today.

This discussion suggests a potential framework for interpreting an important design feature of the GATT/WTO. Under GATT/WTO rules, governments do not negotiate "strong bindings" (i.e., precise tariff levels); instead, they negotiate "weak bindings" (i.e., maximal tariff levels). With an agreement to a use a weak binding, a government thus agrees not to apply its tariff on the relevant product at a level that exceeds the bound rate. The government is free, however, to exercise "downward discretion" and apply a tariff below the bound rate. The preceding discussion suggests that such a design could be attractive to governments as a way of lowering the average tariff while facilitating at least improved downward matching between a government's applied tariff and its political pressure. In fact, many governments often do apply tariffs that are strictly below their negotiated tariff bindings.[3] The discussion suggests as well, however, that circumstances may exist under which governments resist applying tariffs below their bound levels, in order not to be perceived as weak.

In this paper, we pursue these themes at a formal level. We do so by developing a theory of self-enforcing trade agreements among privately informed governments, where the private information that governments possess may be persistent through time. We first put the self-enforcement constraint to the side and develop a number of findings for a static model. We then consider a sequence of dynamic models, in which a government's private information is transitory, perfectly persistent and imperfectly persistent, respectively, in order to better understand the implications of private information for the design of self-enforcing trade agreements.

Our main findings for the static model are as follows. We show that an agreement to use a weak binding offers governments greater expected joint welfare than they can achieve with an agreement to use a strong binding or in the absence of an agreement. We also observe that applied tariffs are often below the bound level, when governments agree to the optimal weak binding. As suggested above, a weak bound on tariffs limits the extent to which tariffs can impose terms-of-trade externalities while at the same time allowing downward discretion so as to facilitate a better match between a government's applied tariff and its level of political pressure. A strong binding, by contrast, limits externalities but eliminates all discretion. Indeed, we show that an agreement to use the optimal strong binding may even offer less expected joint welfare to governments than they achieve in the absence of an agreement. Finally, we characterize the incentive-compatible tariffs that maximize expected joint welfare for governments and show that such second-best tariffs cannot be implemented using weak or strong bindings.

In our analyses of dynamic models, we show that the applied tariffs induced by the optimal weak binding can be achieved as part of a self-enforcing agreement among patient governments

[3]The WTO web site provides statistics on the (applied and bound) tariff profiles of all member governments. See http://www.wto.org/english/news_e/news07_e/tariff_sept07_e.htm

when private information is transitory; however, an enforcement "ratchet effect" arises when political pressures are perfectly persistent, in that a government that exercises downward discretion and applies a tariff below the optimal weak binding reveals that it faces little pressure and thus that it would be unlikely to retaliate in an effective way were its trading partner to cheat on the agreement. In response to this problem, a government with low political pressure may "pool" and apply a tariff at the bound rate and in excess of its optimal level, in order not to be perceived as weak. We also show that the ratchet effect may be overcome when political pressures are imperfectly persistent: governments that are sufficiently patient are then able to achieve the applied tariffs induced by the optimal weak binding as part of a self-enforcing agreement. To achieve these tariffs in a self-enforcing agreement, governments must be more patient when political pressures are more persistent.

For the dynamic model with perfectly peristent types, we also explore an information-based theory of gradualism in trade agreements, in which the tariffs that governments apply under an initial agreement reveal information about the persistent pressures that they face and thus guide the determination of the bound tariff rates that they negotiate in a subsequent agreement. In the equilibrium that we feature, a government with low pressure applies a tariff that is below the bound rate in the initial agreement and then applies a lower tariff that is equal to the (reduced) bound rate in the subsequent agreement. The theory entails "dynamic screening" and suggests that applied tariffs are more likely to be below bound rates in early negotiation rounds. For any positive discount factor that governments may possess and regardless of the pressures that they face, we show that the featured equilibrium generates strictly higher payoffs for governments than they achieve in the absence of an agreement.

This paper contributes to the literature at three broad levels. First, at a methodological level, this paper contributes to the theory of trade agreements among governments with private information. Building on recent work on the theory of collusion among privately informed firms by Athey and Bagwell (2001, 2006), Athey, Bagwell and Sanchirico (2004) and Skryzpacz and Hopenhayn (2004), this paper joins a small and recent literature that explores self-enforcing trade agreements among privately informed governments. This literature includes Bagwell and Staiger (2005), Lee (2007) and Martin and Vergote (2004).[4] With the exception of Athey and Bagwell (2006), all of these papers assume that private information is transitory. When private information exhibits persistence, the analysis becomes more complicated, since a player's behavior may signal his information and thereby affect the beliefs of others players. The present paper appears to be the first analysis of self-enforcing trade agreements between governments with persistent private information.

[4] See also Feenstra and Lewis (1991) and Park (2006). Feenstra and Lewis (1991) consider trade policies among privately informed governments in a static model that allows for transfers between governments. We require here that trade agreements are self-enforcing and do not permit transfers. Park (2006) considers self-enforcing trade agreements among governments that observe private signals about the levels of concealed trade barriers.

Second, at a substantive level, this paper offers new insights with respect to use of bound and applied tariffs in trade agreements. In both the static and dynamic models, we identify circumstances under which governments with low political pressures apply tariffs that are strictly below their bound rates. This finding is also developed by Bagwell and Staiger (2005).[5] The present paper differs on several important dimensions. In the present paper, we extract additional predictions by using a two-type (as opposed to a continuum-type) model. For example, in our static analysis, we present a thorough comparison of tariffs and government welfares when governments use the optimal weak binding, the optimal strong binding, the second-best tariffs and Nash tariffs. In the dynamic model, our two-type framework is sufficiently tractable that we can analyze and compare self-enforcing trade agreements when private information is transitory, perfectly persistent and imperfectly persistent. We show that the self-enforcement constraint requires greater patience from governments when their private information is more persistent. We also identify a new ratchet effect and characterize novel pooling behavior when private information is persistent and weak bindings are used.

Finally, this paper contributes to an established literature that considers explanations for the gradual manner in which trade is liberalized over time through GATT/WTO rounds. For example, Devereux (1997) and Staiger (1995) provide models in which liberalization leads to changes in structural variables (associated with learning by doing or worker locations, respectively), which in turn interact with the self-enforcement constraints in a way that makes further liberalization possible.[6] As discussed above, we present here a new information-based foundation for gradualism in trade agreements

The paper is organized as follows. Section 2 contains the basic model of trade. Our results for the static model are presented in Section 3. In Section 4, we analyze the repeated game with transitory shocks, while Section 5 contains our analysis of the dynamic game with perfectly persistent shocks. The dynamic game with imperfectly peristent shocks is considered in Section 6. In Section 7, we present our results on gradualism. Section 8 concludes, and all remaining proofs are found in the Appendix.

[5]Horn, Maggi and Staiger (2006) develop a related finding in a static model without private information that allows for contracting costs. Maggi and Rodriguez-Clare (2007) analyze a model in which applied tariffs are set at bound levels in equilibrium; however, the potential for governments to apply a tariff below the bound level encourages ex post lobbying and thus mitigates an overinvestment problem. If import-competing firms are risk averse, a government may also keep its applied tariff below the bound rate in order to create policy space, so that the applied tariff can adjust to world-price fluctuations and reduce the variability of the local price. See, for example, Bagwell and Sykes (2004).

[6]For other models of gradualism, see, for example, Bond and Park (2002) and Chisik (2003).

2. The Model

In this section, we present our basic model of trade, and we characterize Nash and efficient tariff choices. We also consider incentive-compatible trade policies.[7]

2.1. Basic Set Up

We consider a partial-equilibrium model, in which trade occurs in two goods between two countries. The home country exports good y and imports good x, while the foreign country exports good x and imports good y. Let P_x and P_y denote the local prices of goods x and y, respectively, in the home country. Using an asterisk to denote foreign country variables, we likewise denote the local prices of goods x and y in the foreign country as P_x^* and P_y^*, respectively. Each good is produced and demanded in each country. For simplicity, we assume that demand functions are symmetric across goods and countries. Gains from trade arise, since the two countries have different domestic supply functions.

We assume that demand and supply functions take simple linear forms. For good $i = x, y$, the home-country demand function is $D(P_i) = 1 - P_i$, and the foreign-country demand function is $D(P_i^*) = 1 - P_i^*$. The home- and foreign-country supply functions for goods x and y are $Q_x(P_x) = P_x/2$, $Q_y(P_y) = P_y$, $Q_x^*(P_x^*) = P_x^*$, and $Q_y^*(P_y^*) = P_y^*/2$, respectively.[8] Notice that the countries are symmetric, in that the supply functions in the foreign country are the mirror image of those in the home country. The corresponding producer surplus (profit) functions are $\pi_x(P_x) = (P_x)^2/4$, $\pi_y(P_y) = (P_y)^2/2$, $\pi_x^*(P_x^*) = (P_x^*)^2/2$ and $\pi_y^*(P_y^*) = (P_y^*)^2/4$. Finally, we may define home- and foreign-country import and export functions as follows: $M_x(P_x) \equiv D(P_x) - Q_x(P_x)$, $E_y(P_y) \equiv Q_y(P_y) - D(P_y)$, $M_y^*(P_y^*) \equiv D(P_y^*) - Q_y^*(P_y^*)$ and $E_x^*(P_x^*) \equiv Q_x^*(P_x^*) - D(P_x^*)$, respectively.

The government of each country selects a specific import tariff. Let τ and τ^* respectively denote the tariffs of the home-country and foreign-country governments. Market-clearing prices are determined once the tariffs are imposed. Home-country exporters of good y receive the world price $P_y^w \equiv P_y$, and similarly foreign-country exporters of good x sell at the world price $P_x^w \equiv P_x^*$. Provided that tariffs do not prohibit trade, the local price for an imported good is determined as $P_x = P_x^w + \tau$ and $P_y^* = P_y^w + \tau^*$. With these relationships in place, we may determine the market-clearing world prices, $P_x^w(\tau)$ and $P_y^w(\tau^*)$, as the world prices that respectively satisfy the following market-clearing conditions: $M_x(P_x^w + \tau) = E_x^*(P_x^w)$ and $M_y^*(P_y^w + \tau^*) = E_y(P_y^w)$. The remaining local prices are then determined as $\widehat{P}_x(\tau) \equiv P_x^w(\tau) + \tau$ and $\widehat{P}_y^*(\tau^*) \equiv P_y^w(\tau^*) + \tau^*$. The explicit solutions are $P_x^w(\tau) = [4 - 3\tau]/7$, $P_y^w(\tau^*) = [4 - 3\tau^*]/7$,

[7] The model that is presented here is the same model that Bagwell and Staiger (2005) consider. This section is thus closely related to the second section of their paper.

[8] Following Bagwell and Staiger (2001), we can derive the supply functions from underlying production functions, in which the home (foreign) country has a higher marginal product of labor for good y (x).

$\widehat{P}_x(\tau) = 4[1+\tau]/7$ and $\widehat{P}^*_y(\tau^*) = 4[1+\tau^*]/7$ for $\tau < 1/6$ and $\tau^* < 1/6$. Trade in a good is prohibited when the import tariff on that good equals or exceeds 1/6.

We now consider government preferences. Following Baldwin (1987), we assume that each government maximizes a weighted sum of consumer surplus, producer surplus and tariff revenue, where due to political-economy pressures each government places a relatively greater weight on producer surplus in its import-competing industry.[9] Formally, let $\gamma \in [1, 7/4]$ denote the weight placed by the government of the home country on the producer surplus enjoyed by its import-competing industry. We may now define the welfare experienced by the government of the home country on its import good as follows:

$$W_x(\widehat{P}_x, P^w_x) = \int_{\widehat{P}_x}^{1} D(P_x)dP_x + \gamma\pi_x(\widehat{P}_x) + [\widehat{P}_x - P^w_x]M_x(\widehat{P}_x).$$

Likewise, on its export good, the government of the home country enjoys the following welfare:

$$W_y(P^w_y) = \int_{P^w_y}^{1} D(P_y)dP_y + \pi_y(P^w_y).$$

The total welfare enjoyed by the government of the home country is then $W_x(\widehat{P}_x, P^w_x) + W_y(P^w_y)$.

The welfare for the government of the foreign country is defined in an analogous fashion. Let $\gamma^* \in [1, 7/4]$ denote the weight placed by the government of the foreign country on the producer surplus in its import-competing industry. The welfares that the government of the foreign country enjoys on its import and export goods are respectively defined as follows:

$$W^*_y(\widehat{P}^*_y, P^w_y) = \int_{\widehat{P}^*_y}^{1} D(P^*_y)dP^*_y + \gamma^*\pi^*_y(\widehat{P}^*_y) + [\widehat{P}^*_y - P^w_y]M^*_y(\widehat{P}^*_y)$$

and

$$W^*_x(P^w_x) = \int_{P^w_x}^{1} D(P^*_x)dP^*_x + \pi^*_x(P^w_x).$$

The total welfare enjoyed by the government of the foreign country is then the sum of these two welfares.

Recalling that local and world prices are determined by tariffs, we observe that government welfares are ultimately functions of tariffs. The home-country tariff, τ, affects the home-country government's welfare on its import good, W_x, and the foreign-country government's welfare on its export good, W^*_x. Thus, we may characterize the Nash and efficient choices of τ with reference only to W_x and W^*_x. Given the symmetry of the model, we henceforth focus on good x. We also observe that only the home-country political-economy parameter, γ, influences the determination of the Nash and efficient choices for τ.

[9] Grossman and Helpman (1994) provide microfoundations for such government preference functions.

In the analysis that follows, it is sometimes convenient to have explicit expressions for W_x and W_x^* as functions of τ and γ. Straightforward calculations yield

$$W_x(\tau;\gamma) = \frac{9+8\gamma}{98} + \frac{8\gamma-5}{49}\tau + \frac{2(2\gamma-17)}{49}\tau^2 \tag{2.1}$$

and

$$W_x^*(\tau) = \frac{25}{98} - \frac{3}{49}\tau + \frac{9}{49}\tau^2, \tag{2.2}$$

where our notation henceforth directly reflects the dependence of welfares on tariffs.

2.2. Nash and Efficient Tariffs

With the welfare functions defined, we may now characterize the Nash and efficient tariffs. We start with the Nash tariff. The optimal unilateral tariff for the government of the home country maximizes $W_x(\tau;\gamma)$. Using (2.1), we find that the first-order condition is satisfied when $\tau = \tau^N(\gamma)$, where

$$\tau^N(\gamma) = \frac{8\gamma-5}{4(17-2\gamma)}. \tag{2.3}$$

For all $\gamma \in [1, 7/4]$, the second-order condition holds; furthermore, τ^N is non-prohibitive for all $\gamma \in [1, 7/4)$, and τ^N reaches the prohibitive level of $1/6$ when $\gamma = 7/4$. Using (2.3), we may confirm that $\tau^N(\gamma)$ is strictly increasing in γ. Intuitively, a higher tariff is more attractive to a government that values more heavily the producer surplus enjoyed in the import-competing industry. The case in which the government of the home country maximizes national income is captured when $\gamma = 1$. In that case, the Nash tariff takes the value $1/20$, which corresponds to the traditional "optimal tariff" for a national-income maximizing government.

The home-country tariff also induces a negative externality on the welfare of the government of the foreign country. Using (2.2), it is straightforward to confirm that $W_x^*(\tau)$ is strictly decreasing in $\tau = \tau^N(\gamma)$ over $\gamma \in [1, 7/4)$. Intuitively, when the government of the home country raises its import tariff, the world price of the foreign country's export good is reduced. This terms-of-trade externality reduces the producer surplus of foreign exporters. Hence, the government of the foreign country is strictly harmed by an increase in the home-country tariff, so long as the foreign country is exporting a positive volume (i.e., provided $\tau < 1/6$). Due to this negative externality, Nash tariffs will not be efficient, when efficiency is measured relative to the joint welfare of the home- and foreign-country governments.

In our partial-equilibrium setting, the efficient tariff choice for the home government is the tariff that maximizes joint welfare on the imported good. Formally, for a given political-economy parameter γ, the efficient tariff is the τ that maximizes $J(\tau;\gamma) \equiv W_x(\tau;\gamma) + W_x^*(\tau)$. Using (2.1) and (2.2), we have that

$$J(\tau;\gamma) = \frac{17+4\gamma}{49} + \frac{8(\gamma-1)}{49}\tau + \frac{4\gamma-25}{49}\tau^2. \tag{2.4}$$

The first order condition for the maximization of (2.4) is satisfied when $\tau = \tau^E(\gamma)$ where

$$\tau^E(\gamma) = \frac{4(\gamma - 1)}{25 - 4\gamma}. \tag{2.5}$$

For all $\gamma \in [1, 7/4]$, the second order condition is satisfied. Notice that the efficient tariff is free trade when the home-country government maximizes national income (i.e., when $\gamma = 1$).

The Nash and efficient tariffs are compared in Figure 1. As illustrated there, for all $\gamma \in [1, 7/4)$, we may use (2.3) and (2.5) to find that $\tau^N(\gamma) > \tau^E(\gamma)$. Intuitively, the Nash tariff exceeds the efficient tariff, since only the latter reflects a concern with the negative externality that flows from the home-country tariff to the foreign-country welfare. At $\gamma = 7/4$, the Nash and efficient tariffs agree, as both then equal 1/6 and thus prohibit all trade. In the analysis below, we assume that the political-economy parameter is either low ($\gamma = L$) or high ($\gamma = H$). To focus attention on political-economy pressures that are never so severe as to result in a prohibition of trade, we assume that $1 \leq L < H < 7/4$.

2.3. Incentive-Compatible Tariffs

We assume that the political-economy parameters, γ and γ^*, are drawn independently from a common distribution. In particular, for any government, the probability that the political-economy parameter takes value L is denoted as η_L where $\eta_L \in (0, 1)$. Likewise, the probability that the political-economy parameter takes value H is $\eta_H \equiv 1 - \eta_L \in (0, 1)$. Each government privately observes its own realized political-economy parameter. Below, we often refer to a government's parameter as its "type."

When governments have private information, some applied tariff schedules fail to be incentive compatible. For example, governments might seek to implement the efficient tariff schedule under which a government applies the tariff that is efficient given its type. Under this arrangement, the government of the home country is expected to apply the tariff $\tau^E(H)$ when its type is H and the tariff $\tau^E(L)$ when its type is L. It may be, though, that $\tau^E(H)$ is closer to $\tau^N(L)$ than is $\tau^E(L)$, so that $W_x(\tau^E(H); L) > W_x(\tau^E(L); L)$. If this is the case, then the home-country government may "lie" and behave as if it has a high type when in fact it has a low type. Intuitively, since the efficient tariff is higher when the home-country government has a high type, this government may be tempted to lie and report a high type as it can then apply a higher tariff and enjoy the consequent terms-of-trade gain.

We next define the incentive-compatibility constraint for the static model. To this end, it is useful to first define $G(\tau) \equiv \frac{9}{98} - \frac{5}{49}\tau - \frac{34}{49}\tau^2$ and $f(\tau) \equiv \frac{4}{49} + \frac{8}{49}\tau + \frac{4}{49}\tau^2$. Note that $f(\tau) > 0$ and $f'(\tau) > 0$ for all $\tau \in [0, 1/6]$. Using (2.1), we see that $W_x(\tau; \gamma) = G(\tau) + \gamma f(\tau)$. Suppose now that governments seek to implement an applied tariff schedule, $\tau(\gamma)$. Given this schedule, we may define the home-country government welfare for good x that is enjoyed when this government's actual type is γ and it applies the tariff, $\tau(\widehat{\gamma})$, that is intended for type $\widehat{\gamma}$:

$$\widetilde{W}_x(\widehat{\gamma}, \gamma) \equiv G(\tau(\widehat{\gamma})) + \gamma f(\tau(\widehat{\gamma})). \tag{2.6}$$

The applied tariff schedule $\tau(\gamma)$ is *incentive compatible* if and only if, for all $\gamma \in \{L, H\}$ and $\widehat{\gamma} \in \{L, H\}$,

$$\widetilde{W}_x(\gamma, \gamma) \geq \widetilde{W}_x(\widehat{\gamma}, \gamma). \tag{2.7}$$

Using standard arguments, we may confirm that an applied tariff schedule, $\tau(\gamma)$, is incentive compatible only if it is nondecreasing.[10]

3. Static Model: Analysis

In this section, we analyze the static model. We distinguish between two kinds of bounds on applied tariffs. In an agreement with *strong bindings*, a government must apply a tariff that equals the bound rate; whereas, in an agreement with *weak bindings*, a government can apply any tariff that does not exceed the bound rate. We characterize the optimal strong and weak bindings, and we compare expected government welfare under these bindings with the expected welfare achieved under Nash tariffs. We also characterize the most efficient tariffs that are incentive compatible.

3.1. Optimal Strong Bindings

Under an agreement with strong bindings, each government agrees to set its applied tariff precisely at the bound rate, regardless of the political pressure that it faces. When a strong binding is in place, a government thus does not have discretion to lower or raise its tariff from the bound rate.

A strong binding has benefits and costs. On the one hand, due to the terms-of-trade externality, Nash tariffs are higher than is efficient, and a strong binding can offer the benefit of preventing high tariffs. Indeed, in an environment with complete information, the strong binding could be set equal to the efficient tariff, and a strong binding would then deliver an efficient outcome. On the other hand, when we allow for incomplete information, a strong binding has a potential cost, in that it eliminates a government's ability to exercise discretion and set its applied tariff at a rate that reflects its political situation. Thus, an agreement with a strong binding cannot deliver efficient tariffs when private information is present. A more subtle issue is whether an agreement with a strong binding can always be found which offers greater expected welfare to governments than they would be able to achieve in the absence of an agreement (i.e., in the Nash equilibrium). We address this issue below.

Let τ^S denote a strong binding. In an agreement with a strong binding, the induced applied tariff schedule is $\tau(\gamma) \equiv \tau^S$. If governments negotiate an agreement with a strong binding, then

[10] Use (2.6) and observe that incentive compatibility requires $\widetilde{W}_x(\gamma, \gamma) \geq \widetilde{W}_x(\widehat{\gamma}, \gamma) = \widetilde{W}_x(\widehat{\gamma}, \widehat{\gamma}) + (\gamma - \widehat{\gamma}) f(\tau(\widehat{\gamma}))$ and $\widetilde{W}_x(\widehat{\gamma}, \widehat{\gamma}) \geq \widetilde{W}_x(\gamma, \widehat{\gamma}) = \widetilde{W}_x(\gamma, \gamma) + (\widehat{\gamma} - \gamma) f(\tau(\gamma))$. It follows that $(\gamma - \widehat{\gamma}) f(\tau(\gamma)) \geq \widetilde{W}_x(\gamma, \gamma) - \widetilde{W}_x(\widehat{\gamma}, \widehat{\gamma}) \geq (\gamma - \widehat{\gamma}) f(\tau(\widehat{\gamma}))$. Thus, $f(\tau(\gamma))$ must be nondecreasing in γ. Given $f' > 0$, an applied tariff schedule, $\tau(\gamma)$, is incentive compatible only if it is nondecreasing.

they may seek the strong binding that maximizes joint expected welfare. We thus define an *optimal strong binding*, τ_A^S, as the solution to the following program:

$$\max_{\tau^S} EJ(\tau^S;\gamma),$$

where E is the expectation operation taken with respect to γ. We do not include here an incentive-compatibility constraint, since the applied tariffs induced by a strong binding do not vary with type and thus are trivially incentive compatible. Given that $W_x(\tau;\gamma)$ is linear in γ, the optimal strong binding must maximize $J(\tau^S;E\gamma)$, where $E\gamma \equiv \eta_L L + \eta_H H$. Using (2.5), we thus conclude that

$$\tau_A^S = \tau^E(E\gamma) = \frac{4(E\gamma - 1)}{25 - 4E\gamma}. \tag{3.1}$$

In short, the optimal strong binding is the efficient tariff for the expected type.

We seek to compare expected joint welfare when governments reach an agreement to use the optimal strong binding with the expected joint welfare when governments have no agreement and thus apply Nash tariffs. For a given good, it is clear that the expected welfare for the government of the importing country (i.e., EW_x) is higher when it is allowed to apply Nash tariffs. It is more interesting to consider expected welfare for the government of the exporting country (i.e., EW_x^*). As suggested above, it reasonable to anticipate that an agreement to use the optimal strong binding would lower the expected applied tariff. Such an agreement would also reduce the variance in the applied tariff. Since the welfare of the government of the exporting country is decreasing and convex in the applied tariff, it is not immediately clear whether an agreement to use an optimal strong tariff would increase the expected welfare of the exporting country.

Despite these conflicting considerations, we are able to report the following:

Proposition 1: *(i) The expected Nash tariff is strictly higher than the optimal strong binding: $E\tau^N(\gamma) > \tau_A^S$. (ii) Expected welfare for the government of the exporting country is strictly higher when the applied tariff is induced by the optimal strong binding than when the Nash tariffs are applied: $W_x^*(\tau_A^S) > EW_x^*(\tau^N(\gamma))$.*

Thus, for a given good, the government of the exporting country strictly prefers the applied tariff induced by the optimal strong binding to the applied Nash tariffs, while the government of the importing country has the opposite preference. The proof is found in the Appendix.

We next analyze the effect of an agreement to use the optimal strong binding on expected joint welfare. Before undertaking this analysis, however, we impose a parameter restriction that is maintained throughout the remainder of the paper. In particular, we assume that the optimal strong binding exceeds the Nash tariff when the type is low: $\tau_A^S = \tau^E(E\gamma) > \tau^N(L)$. As we confirm below, when this inequality holds, the optimal weak and strong bindings induce different applied tariffs, and the optimal weak binding is such that a government with a low type applies a tariff that is strictly below the bound level. Using (2.3) and (3.1), we may restate the desired inequality in terms of the following restriction on parameters: $E\gamma > (8L + 7)/12$.

Consider now Figure 2. As shown there, the optimal strong binding induces an applied tariff that is higher (lower) than would be efficient when the type is low (high). Given our parameter restriction, the optimal strong binding induces an applied tariff that is higher than the applied Nash tariff when the type is low. Thus, in the event that the government of the importing country has a low type, the optimal strong binding is less efficient than is the corresponding Nash tariff. By contrast, in the event that the type is high, the efficient tariff is below the corresponding Nash tariff and above the applied tariff that is induced by the optimal strong binding. A trade-off may thus arise across the two events, since the optimal strong binding may be more efficient than the Nash tariff when the type is high.

Looking at Figure 2, however, we may anticipate one circumstance in which a trade-off across events fails to arise. In particular, suppose that the difference in types is large, so that H is near $7/4$ (i.e., the upper bound at which the Nash tariff prohibits trade) while L is well below this bound. The difference between $\tau^N(H)$ and $\tau^E(H)$ is then small, and so the Nash tariff is approximately efficient when the type is high. The optimal strong binding, however, equals $\tau^E(E\gamma)$ and thus remains below $\tau^E(H)$, with the gap being larger when η_H is smaller. Thus, when $H - L$ is large in this sense and η_H is not too high, we anticipate that joint welfare is higher when Nash tariffs are applied than when an agreement to use the optimal strong binding is reached, whether the importing government's type is low or high. If this reasoning is correct, then circumstances exist under which an agreement to use a strong binding cannot improve upon the no-agreement benchmark in which Nash tariffs are applied.

We introduce a proposition below that confirms this reasoning and offers further comparisons. To state the proposition, we require a few definitions. Consider first our parameter restriction that $E\gamma > (8L+7)/12$. This restriction can be expressed equivalently as $\eta_H > \frac{1}{3}\frac{7/4-L}{H-L}$ and thus implies $\eta_H > 1/3$. In fact, our restriction can be restated in equivalent form as the requirements that $\eta_H > 1/3$ and

$$L(H, \eta_H) \equiv (3\eta_H H - 7/4)/(3\eta_H - 1) > L. \tag{3.2}$$

For a fixed $\eta_H > 1/3$, and given $H < 7/4$, we see that $L(H, \eta_H)$ is strictly increasing in H and η_H and that $L(H, \eta_H) < H$. It is also convenient to note that $L(7/4, \eta_H) \equiv 7/4$.

Second, we set $J(\tau^N(L), H) = J(\tau^N(H), H)$ and thereby implicitly define the function $L = L^N(H)$. This function defines the combinations of L and H such that joint welfare when the government of the importing country has a high type is the same whether the high or low Nash tariff is applied. Put differently, this function describes combinations of L and H such that $\tau^E(H) = (\tau^N(H) + \tau^N(L))/2$. We find that

$$L^N(H) = (143H - 119 - 12H^2)/(61 - 4H). \tag{3.3}$$

Using (3.2) and (3.3), calculations confirm the following relationships: $L^N(H)$ is strictly increasing in H, $L^N(7/4) = 7/4$, and $L(H,1) > L^N(H) > L(H, 1/3+\varepsilon)$ for $\varepsilon > 0$ and sufficiently small. We note as well that

$$sign\{J(\tau^N(L); H) - J(\tau^N(H); H)\} = sign\{L - L^N(H)\}. \tag{3.4}$$

The relationship between $L(H, \eta_H)$ and $L^N(H)$ is depicted in Figure 3.

We are now ready to state our second proposition. The proposition concerns $\Delta \equiv EJ(\tau_A^S; \gamma) - EJ(\tau^N(\gamma); \gamma)$, which is the difference in the expected joint welfare achieved under an agreement to use the optimal strong binding and that achieved in the Nash equilibrium.

Proposition 2: *(i). Suppose $H - L$ is small in the sense that $L(H,1) > L > L^N(H)$. There exists $\widetilde{\eta}_H \in (1/3, 1)$ such that for all $\eta_H \in (\widetilde{\eta}_H, 1)$, $\tau_A^S > \tau^N(L)$ and $\Delta > 0$. (ii). Suppose $H - L$ is large in the sense that $L^N(H) > L$. There exists $\widetilde{\eta}_H \in (1/3, 1)$ and $\widehat{\eta}_H \in (\widetilde{\eta}_H, 1)$ such that (a). for all $\eta_H \in (\widetilde{\eta}_H, 1)$, $\tau_A^S > \tau^N(L)$, (b). for all $\eta_H \in (\widetilde{\eta}_H, \widehat{\eta}_H)$, $\Delta < 0$, and (c). for all $\eta_H \in (\widehat{\eta}_H, 1)$, $\Delta > 0$.*

The proof of this proposition is found in the Appendix.

Proposition 2 may be understood with reference to Figure 3. Part (i) of the proposition applies when L and H reside in region A. For a particular point $x = (H, L)$ in this region, there exists a critical value $\widetilde{\eta}_H \in (1/3, 1)$ such that $L = L(H, \widetilde{\eta}_H)$. In the figure, the parameter value $\widetilde{\eta}_H$ generates the dotted line that passes through x. For any $\eta_H \in (\widetilde{\eta}_H, 1)$, the point x lies below $L(H, \eta_H)$. Our parameter restriction then holds and thus $\tau_A^S > \tau^N(L)$. Part (i) of the proposition indicates that the optimal strong binding then generates strictly higher expected welfare than occurs in the Nash equilibrium. Intuitively, when $H - L$ is small, the volume of trade is large even when the government of the importing country has a high type and sets its Nash tariff. The presence of a large trade volume in turn implies that the terms-of-trade externality plays an important role, with the result being a large difference between the Nash and efficient tariffs in the high-type state.

Part (ii) of the proposition concerns region B. For a particular point $y = (L, H)$ in this region, we may again find $\widetilde{\eta}_H \in (1/3, 1)$ such that $L = L(H, \widetilde{\eta}_H)$. For η_H slightly above this value, τ_A^S is only slightly above $\tau^N(L)$. Joint welfare is then approximately the same in the low-type state whether governments have an agreement or not. In this region, however, the volume of trade is low when the government has a high type and applies its Nash tariff. Accordingly, the terms-of-trade externality then plays a small role, and so the Nash tariff is close to efficient in the high-type state. Expected joint welfare is then strictly higher for governments when they have no agreement and thus apply their Nash tariffs. Finally, for higher values of η_H, an agreement to use the optimal strong binding generates strictly higher expected joint welfare. To see this, note that when $\eta_H \approx 1$, so that the government of the importing country almost always has a high type, the optimal strong binding induces an applied tariff that is approximately equal to $\tau^E(H)$. An agreement to use the optimal strong binding then induces an applied tariff that is almost always approximately efficient.

In general, when governments are privately informed as to their political pressures, a trade agreement has two roles to serve. First, it should reduce the average tariff and thereby undo the inefficiency that is attributable to the terms-of-trade externality. Second, it should provide sufficient flexibility that governments can adjust their tariffs to their political circumstances. An agreement to use an optimal strong binding addresses the first role and neglects the second. As Proposition 2 indicates, in some cases, such an agreement generates *lower* expected joint welfare for governments than they achieve in the absence of any agreement. Proposition 2 thus provides motivation to consider alternative forms for agreements, in which the average tariff is lowered and yet governments are also able to exercise some discretion in an incentive-compatible way. We thus next consider an agreement to use a weak binding.

3.2. Optimal Weak Bindings

We consider now an agreement to use a weak binding. With such an agreement, governments place an upper bound on the permissible level of an applied tariff, and at the same time provide downward discretion so that a government can apply a tariff below the binding if it so chooses. An agreement with a weak binding thus offers scope for lowering the average tariff from the Nash level while also providing downward flexibility.

Let τ^W denote a weak binding. In an agreement with a weak binding, the induced applied tariff schedule is $\tau(\gamma) = \min\{\tau^W, \tau^N(\gamma)\}$. Intuitively, a government applies a tariff below the bound rate, if its optimal tariff is below the bound rate. Otherwise, the government applies a tariff equal to the bound rate, as it thereby applies a tariff that is as close to its optimal rate as possible. When governments negotiate an agreement with a weak binding, they may select the weak binding that maximizes expected joint welfare. Accordingly, we define an *optimal weak binding*, τ_A^W, as the solution to the following program:

$$\max_{\tau^W} EJ(\min\{\tau^W, \tau^N(\gamma)\}; \gamma),$$

where the expectation is taken over γ. It is not necessary to include an incentive compatibility constraint, since this constraint is automatically satisfied: a weak binding precludes (places no restrictions on) the application of a tariff above (below) the bound rate.

Consider now various candidates for τ^W. First, if $\tau^W \geq \tau^N(H)$, then a government would set its import tariff at $\tau^N(H)$ when its type is high and at $\tau^N(L)$ when its type is low. Thus, a weak binding in this range simply induces the application of Nash tariffs. We may conclude that, when governments agree on the optimal weak binding, their expected joint welfare cannot be lower than that which they achieve under Nash tariffs. Recall that the expected joint welfare under Nash tariffs sometimes exceeds that which arises when governments use the optimal strong binding. Second, at the other extreme, if $\tau^W \leq \tau^N(L)$, then the weak binding is never above a government's optimal tariff, and so governments would always apply a tariff that is equal to the weak binding. In this case, the weak binding performs like a strong binding, since no downward discretion is excercised when the weak binding is set at such a low level. Under our parameter restriction, as argued above, the optimal strong binding is $\tau_A^S = \tau^E(E\gamma) > \tau^N(L)$. Given that expected joint welfare is concave in the applied tariff, it follows that the optimal weak binding must satisfy $\tau_A^W \geq \tau^N(L)$.

The third possibility is that $\tau^W \in (\tau^N(L), \tau^N(H))$. Suppose for example that the weak binding is set at the level of the optimal strong binding: $\tau^W = \tau_A^S$. Given that $\tau^N(H) > \tau_A^S > \tau^N(L)$, if this weak binding is used and a government has a high type, then it would apply a tariff equal to the binding. This weak binding thus induces the same applied tariff for a high type as does the optimal strong binding. Now consider a government with a low type. Since the binding is weak, this government would venture below the binding and apply its optimal tariff, $\tau^N(L)$. Notice, too, that joint welfare is then strictly increased, since $\tau^N(L)$ is closer to $\tau^E(L)$ than is τ_A^S. Since a weak binding with $\tau^W = \tau_A^S$ thus generates strictly higher expected welfare than does the optimal strong binding, we may conclude that governments are sure to achieve strictly higher expected joint welfare when they agree on an optimal weak binding than when they agree on an optimal strong binding. Intuitively, a weak binding can bound the maximal tariff just as effectively as does a strong binding; however, a weak binding offers the further advantage of providing efficiency-enhancing downward discretion.

Our final task is to determine the optimal weak binding. It is now clear that $\tau_A^W > \tau^N(L)$. A weak binding that is at or below $\tau^N(L)$ has the same effect as a strong binding at the same level; however, the optimal strong binding is τ_A^S, and we just argued that a weak binding with $\tau^W = \tau_A^S$ would offer strictly higher expected joint welfare than would a strong binding at this level. For any $\tau^W > \tau^N(L)$, a government with a low type applies its optimal tariff, $\tau^N(L)$. Thus, an optimal weak binding only affects the tariff that is applied by a government with a high type. Consequently, the optimal weak binding is set at the efficient tariff for a government with a high type: $\tau_A^W = \tau^E(H)$.

Observe that the optimal weak binding is strictly higher than the optimal strong binding: $\tau_A^W = \tau^E(H) > \tau^E(E\gamma) = \tau_A^S$. Intuitively, the optimal strong binding induces an applied tariff that is below the level that would be efficient when the government has a high type, since under a strong binding the applied tariff for a high type cannot be increased without increasing, too, the applied tariff for a low type. With a weak binding, governments have more flexibilitiy, and it becomes possible to induce a higher and more efficient tariff for the high type of government without altering the tariff that a low type would apply.

We can summarize our results to this point with the following proposition.

Proposition 3: *When governments reach an agreement to use an optimal weak binding, they enjoy expected joint welfare that is strictly higher than that which they would achieve in an agreement to use an optimal strong binding and than that which they would achieve if Nash* *tariffs were applied. The optimal weak binding is strictly higher than the optimal strong binding. When governments agree to use the optimal weak binding, a government with a low level of political pressure applies its optimal tariff, and this tariff is strictly below the bound rate.*

This result provides one interpretation for the observation that governments often apply tariffs that are below the weak bound rates to which they agree in GATT/WTO negotiations. Bagwell and Staiger (2005) report a similar result in a model with a continuum of types. The two-type model considered here allows for a simple proof of the result. As we show below, in the two-type setting, we can also report new results about the magnitude of expected applied tariffs. Finally, we note that Proposition 3 also leads to the prediction that a government's applied tariff should be higher when the export supply function is less elastic, in the event that the applied tariff is below the bound rate. This follows since the applied tariff is then an optimal tariff (given the government's preferences).[11]

Proposition 3 establishes that the optimal weak binding exceeds the optimal strong binding. An interesting question is whether the applied tariffs that are induced by the respective bindings can be similarly ranked. We next show that, given our parameter restriction, the optimal weak binding induces a strictly higher expected applied tariff than does the optimal strong binding.[12] This is perhaps somewhat surprising, since the traditional terms-of-trade reasoning suggests that expected joint welfare gains are achieved through reductions in tariffs. Intuitively, when governments use a strong binding, they are unable to customize their applied tariffs to their political circumstances; hence, they simply focus on reducing the average applied tariff. When weak bindings are used, however, governments have some flexibility, and they are able to raise the applied tariff to a more efficient level when political pressures are high.

Proposition 4: *The expected Nash tariff is strictly higher than the expected applied tariff when governments agree to use the optimal weak binding, and the expected applied tariff when governments agree to use the optimal weak binding is strictly higher than the optimal strong binding.*

The proof is found in the Appendix.

3.3. Second-Best Tariffs

In this subsection, we consider all of the applied tariff schedules that are incentive compatible when transfers are not feasible. Our analysis thus includes the applied tariffs that are induced by strong and weak bindings; however, we no longer require that the applied tariff schedule

[11] See Broda, Limao and Weinstein (forthcoming) for an examination of the empirical relationship between export supply elasticities and the import tariffs applied by governments prior to accession in the WTO. Their analysis provides empirical support for the hypothesis that governments set optimal tariffs prior to accession.

[12] Without our parameter restriction, it is possible that the optimal weak and strong bindings would be set at a common level below $\tau^N(L)$ and thus induce the same applied tariff.

can be induced by a strong or weak binding. Our findings in this subsection thus provide a benchmark against which an agreement with strong or weak bindings might be compared.

Let $\tau(\gamma)$ denote an incentive-compatible applied tariff schedule. We define the *second-best tariffs* as the applied tariff schedule, $\tau^{SB}(\gamma)$, that solves the following *second-best program*:

$$\max_{\tau(\gamma)} EJ(\tau(\gamma);\gamma) \text{ s.t. } \tau(\gamma) \text{ satisfies (2.7)}$$

We establish above that an applied tariff schedule is incentive compatible only if it is nondecreasing. Our goal now is to characterize the second-best tariffs.

We distinguish between two cases. In the first case, we have that

$$W_x(\tau^E(L);L) \geq W_x(\tau^E(H);L). \tag{3.5}$$

When this first case holds, the second-best tariffs are in fact first best. In other words, when this case holds, the efficient tariff schedule, $\tau^E(\gamma)$, satisfies (2.7) and thus solves the second-best program. To confirm that $\tau^E(\gamma)$ satisfies (2.7), we simply note that a government with a low type prefers applying $\tau^E(L)$ to $\tau^E(H)$ under (3.5). Further, a government with a high type prefers applying $\tau^E(H)$ to $\tau^E(L)$, since $\tau^E(H)$ is closer to this government's optimal tariff.

We find that (3.5) holds if and only if $(\tau^E(H)+\tau^E(L))/2 \geq \tau^N(L)$, so that $\tau^E(L)$ is closer to $\tau^N(L)$ than is $\tau^E(H)$. In turn, this inequality holds if and L and H are such that $H \geq H^{FB}(L) \equiv (9L - 4L^2 + \frac{175}{4})/(41-8L)$. Let $L^{FB}(H)$ denote the inverse of $H^{FB}(L)$, for the values that L and H may take in our model. Then (3.5) holds if and only if $L^{FB}(H) \geq L$, where $L^{FB}(65/44) = 1$ and $L^{FB}(H)$ is strictly increasing and convex in H for $H \in [65/44, 7/4)$. We note that $L^{FB}(7/4) = 7/4$. As shown in Figure 4, $L^{FB}(H)$ lies strictly below $L(H,1)$ and strictly above $L(H,1/2)$. The region for which our first case holds is the set of L and H that lie on and to the southeast of $L^{FB}(H)$. Over this region, the difference between H and L is relatively large. We note as well that, if $\eta_H \in (1/3, 1/2]$, then our parameter restriction (i.e., L and H lie on or to the southeast of $L(H,1/2)$) implies that (3.5) holds.

We summarize our findings with respect to the first case as follows:

Proposition 5: *When (3.5) holds, the efficient tariffs are incentive compatible, and thus $\tau^{SB}(\gamma) = \tau^E(\gamma)$. In turn, (3.5) holds if $H - L$ is large, in the sense that L and H lie on or to the southeast of $L^{FB}(H)$. Under our parameter restriction, if $\eta_H \in (1/3, 1/2]$, then (3.5) must hold.*

Thus, in our two-type model, if the types are far apart, efficient tariffs are incentive compatible.

Recall that the optimal weak binding does not induce the application of the efficient tariffs. It is instructive now to review the logic behind this finding. If a weak binding were to induce the application of efficient tariffs, the bound rate would need to be at least $\tau^E(H)$. Given that $\tau^E(H) > \tau^E(E\gamma) > \tau^N(L)$ under our parameter restriction, it would then follow that the bound rate must strictly exceed $\tau^N(L)$. But then a government with a low type would apply $\tau^N(L)$ rather than $\tau^E(L)$. Hence, it is impossible to induce the application of efficient tariffs by using a weak binding.[13] While the applied tariffs induced by the optimal weak binding are more efficient than those induced by the optimal strong binding and than the Nash tariffs, they do not achieve full efficiency.

The second case holds when (3.5) fails. Thus, in the second case,

$$W_x(\tau^E(L); L) < W_x(\tau^E(H); L). \tag{3.6}$$

In terms of Figure 4, the region for which (3.6) holds is the set of L and H that lie to the northwest of $L^{FB}(H)$. Clearly, efficient tariffs fail to be incentive compatible in this case: a government with a low type would "lie" and pretend to have a high type. In this second case, the second-best tariffs must involve a distortion from the efficient tariffs, for at least one or possibly both types.

To begin our analysis of the second case, we define a *relaxed program*:

$$\max_{\tau(\gamma)} EJ(\tau(\gamma); \gamma) \text{ s.t. } \tau(H) \geq \tau(L) \text{ and } W_x(\tau(L); L) \geq W_x(\tau(H); L).$$

To see that this program is indeed a relaxation of the second-best program, recall that (2.7) implies that applied tariff schedule is nondecreasing and note that the relaxed program omits the constraint that $W_x(\tau(H); H) \geq W_x(\tau(L); H)$. We now solve the relaxed program by establishing three claims. After solving the relaxed program, we verify in the proof to Proposition 6 that the omitted constraint is automatically satisfied and thus that we have solved the second-best program as well.

Our first claim is as follows: When (3.6) holds, at any solution $\tau(\gamma)$ to the relaxed program, $W_x(\tau(L); L) = W_x(\tau(H); L)$. To establish this claim, assume (3.6) holds and suppose that $W_x(\tau(L); L) > W_x(\tau(H); L)$ at a solution to the relaxed program. Under this supposition, $\tau(H) > \tau(L)$ is necessary. Now, if $\tau(L) \neq \tau^E(L)$, then we could move the applied tariff for the low type slightly towards $\tau^E(L)$ and achieve a higher value for expected joint welfare, while still satisfying all of the constraints of the relaxed program. Thus, $\tau(L) = \tau^E(L)$ is necessary. By a similar logic, if $\tau(H) \neq \tau^E(H)$, then we could move the applied tariff for the high type slightly in the direction of $\tau^E(H)$ and achieve a higher value for expected joint welfare, while still satisfying all of the constraints of the relaxed program. Thus, under our supposition, a solution to the relaxed program must satisfy $\tau(L) = \tau^E(L)$ and $\tau(H) = \tau^E(H)$. But this leads to a contradition, since the efficient tariff schedule fails to satisfy the relaxed program's incentive-compatibility constraint under our assumption that (3.6) holds.

[13] In fact, whether or not our parameter restriction holds, it is impossible to induce the application of efficient tariffs using a weak binding. Given $\tau^E(H) < \tau^N(H)$, a government with the high type applies $\tau^E(H)$ only if the weak binding is set at $\tau^E(H)$. If $\tau^E(H) \geq \tau^N(L)$ ($\tau^E(H) < \tau^N(L)$), then a government with a low type applies $\tau^N(L)$ ($\tau^E(H)$). Recall now that $\tau^E(L) < \tau^N(L)$ and $\tau^E(L) < \tau^E(H)$.

Our second claim offers a slight strengthening of the first claim: When (3.6) holds, at any solution $\tau(\gamma)$ to the relaxed program, $W_x(\tau(L);L) = W_x(\tau(H);L)$ and $\tau(H) > \tau(L)$. To establish this claim, we need only show that a solution to the relaxed program cannot set $\tau(H) = \tau(L)$. The constant tariff schedule that maximizes expected joint welfare is derived above in our discussion of the optimal strong binding and entails $\tau(H) = \tau(L) = \tau^E(E\gamma)$. As established in Proposition 3, expected joint welfare would be higher if governments set $\tau(H) = \tau^E(H) > \tau^N(L) = \tau(L)$, where the inequality follows from our parameter restriction. This is the applied tariff schedule that is induced by the optimal weak binding, and these applied tariffs satisfy the constraints of the relaxed program.

We come now to our third claim: When (3.6) holds, at any solution $\tau(\gamma)$ to the relaxed program, $\tau^E(L) < \tau(L) < \tau^N(L)$ and $\tau^E(H) < \tau(H)$. To establish this claim, we begin by using our second claim to see that a solution must satisfy $W_x(\tau(L);L) = W_x(\tau(H);L)$ and $\tau(H) > \tau^N(L) > \tau(L)$. Thus, $\tau(H)$ must be the "flipside" of $\tau(L)$ along the function $W_x(\tau;L)$. As illustrated in Figure 5, let $\overline{\tau}^E(L)$ be the flipside of $\tau^E(L)$. Thus, $W_x(\tau^E(L);L) = W_x(\overline{\tau}^E(L);L)$ and $\overline{\tau}^E(L) > \tau^N(L) > \tau^E(L)$. When (3.6) holds, $W_x(\overline{\tau}^E(L);L) < W_x(\tau^E(H);L)$, and so $\overline{\tau}^E(L) > \tau^E(H) > \tau^N(L)$, where the last inequality follows from our parameter restriction.

When (3.6) holds, we now establish that a solution to the relaxed program must satisfy $\tau(L) > \tau^E(L)$. Suppose first that a solution entails $\tau(L) < \tau^E(L)$. Then $\tau(H)$ must be the flipside of $\tau(L)$. This case corresponds to the points denoted by "1" in Figure 5, and we see that $\tau(H) > \overline{\tau}^E(L) > \tau^E(H)$ follows. Starting here, expected joint welfare could be strictly increased by raising the tariff assigned to the low type while likewise slightly decreasing the flipside tariff that is assigned to the high type. For small movements of this kind, both tariffs move toward their efficient levels, and so expected joint welfare is strictly higher. Since these adjusted tariffs also satisfy the constraints of the relaxed program, we conclude that $\tau(L) \geq \tau^E(L)$. Suppose second that a solution entails $\tau(L) = \tau^E(L)$ and thus $\tau(H) = \overline{\tau}^E(L)$. This case corresponds to the points marked "2" in Figure 5. Now consider again a slight increase in the tariff assigned to the low type with a corresponding slight decrease in the flipside tariff that is assigned to the high type. For a small change, the change in the low-type tariff has no first-order effect on expected joint welfare, since the tariff for the low type is initially set at the efficient level for this type. The change in the tariff for the high type, however, generates a first-order increase in expected welfare, since the tariff for the high type is initially above the level that is efficient for this type. We thus conclude that $\tau(L) > \tau^E(L)$.

When (3.6) holds, we next establish that a solution to the relaxed program must satisfy $\tau(H) > \tau^E(H)$. Suppose first that $\tau^N(L) < \tau(H) < \tau^E(H)$. Let $\tau(L)$ be the (lower) flipside of $\tau(H)$, in that $\tau(L) < \tau(H)$ and $W_x(\tau(L);L) = W_x(\tau(H);L)$. This case corresponds to the points marked "3" in Figure 5. From here, we may consider a slight increase in the tariff for the high type that is accompanied by a flipside decrease in the tariff for the low type. The modified tariffs continue to satisfy the constraints of the relaxed program. They also generate strictly higher expected joint welfare, since for each type the tariff is moved in the direction of the tariff that is efficient for that type. We conclude that $\tau(H) \geq \tau^E(H)$. Suppose second that a solution entails $\tau(H) = \tau^E(H)$ and thus $\tau(L) = \underline{\tau}^E(H)$ where as illustrated $\underline{\tau}^E(H)$ is the (lower) flipside of $\tau^E(H)$. This case corresponds to the points marked "4" in Figure 5. From here, we may increase slightly the tariff for the high type while likewise decreasing slightly the (lower) flipside tariff for the low type. The former adjustment has no first-order effect on expected joint welfare, since the tariff for the high type is initially set at the efficient level for this type. The latter adjustment, however, generates a first-order increase in expected joint welfare, since the tariff for the low type is initially above the level that is efficient for this type. We thus conclude that $\tau(H) > \tau^E(H)$. This completes the proof of our third claim.

We are now ready to state and prove the following propostion:

Proposition 6: *When (3.6) holds, $\tau^{SB}(\gamma)$ satisfies $\tau^E(L) < \tau^{SB}(L) < \tau^N(L)$ and $\tau^E(H) < \tau^{SB}(H)$. In turn, (3.6) holds if $H - L$ is small, in the sense that L and H lie on or to the northwest of $L^{FB}(H)$. There exists values for L and H that satisfy (3.6) and our parameter restriction, if η_H is sufficiently large.*

The remainder of the proof is found in the Appendix. We establish there that the solution to the relaxed program satisfies the omitted incentive constraint and thus is also the solution to the second-best program. We also confirm there the stated relationships between (3.6) and the parameters of the model.

It is interesting to compare the second-best tariffs characterized in Proposition 6 with the applied tariffs induced by the optimal weak binding. We make two observations. First, like the efficient tariffs, the second-best tariffs characterized in Proposition 6 cannot be induced through an agreement to use a weak (or strong) binding. Instead, an agreement that implements the second-best tariffs would allow two particular tariffs, $\tau(L)$ and $\tau(H)$, and regard any other selection as a violation. An agreement of this kind requires careful calibration and monitoring; thus, it is perhaps understandable that this form of agreement is not descriptive of actual GATT/WTO tariff agreements. Nevertheless, it is instructive to understand the form that an optimal agreement takes, if governments possess private information and are otherwise free from transaction costs. Second, the second-best tariffs entail a wider spread than do the applied tariffs induced by the optimal weak binding; in particular, under the second-best tariff schedule, the lowest tariff is lower ($\tau^{SB}(L) < \tau^N(L)$) and the highest tariff is higher ($\tau^{SB}(H) > \tau^E(H)$).

It is also interesting to compare the second-best tariffs characterized in Proposition 6 with the efficient tariffs. Notice that the second-best tariffs entail distortions at the bottom and at the top: the low type of government applies a tariff that is above the efficient level ($\tau^{SB}(L) > \tau^E(L)$), and the high type of government also applies a tariff that exceeds the efficient level ($\tau^{SB}(H) > \tau^E(H)$). This finding contrasts with the usual "no-distortion-at-the-top" finding that obtains in a standard Principal-Agent problem, in which a Principal maximizes its objective and provides a menu of choices to a privately informed Agent. The problem we analyze here is different in two important respects. First, we do not have a Principal; instead, we maximize Agents' expected joint welfare and provide a menu of choices to privately informed Agents. Second, unlike the standard Principal-Agent problem, our analysis does not allow for transfers. The absence of transfers makes the second-best program more challenging to analyze. When transfers are allowed, a straightforward argument indicates that governments can design a trade agreement in which they achieve efficient tariffs.[14]

Our analysis here is also similar in some ways to that in recent work on repeated games with private information by Athey, Atkeson and Kehoe (2005), Athey and Bagwell (2001, 2006) and Athey, Bagwell and Sanchirico (2004). In the stage games emphasized in this work, one player's expected welfare is affected by the actions of another player through the probability that the latter player's action is higher or lower than the former player's action or through the mean action level of the latter player. Correspondingly, under appropriate distributional assumptions, the second-best schedule for the Agents involves rigid behavior (i.e., a strong binding) or downward discretion with a cap (i.e., a weak binding). One novel feature of our analysis is that government welfare functions are quadratic in tariffs. As (2.2) confirms, one government's welfare is affected by the mean and variance of the other government's tariff. Working with a two-type model, we offer in Proposition 6 a simple characterization of the second-best tariffs for this linear-quadratic setting. Interestingly, the second-best tariffs cannot be achieved with a strong or weak binding.

4. Enforcement: Transitory Shocks

In the previous section, we analyze a static model and simply assume that any agreement can be enforced. In fact, trade agreements must be self-enforcing. In the remainder of paper, we therefore focus on dynamic models. This enables us to formalize the idea that a government will honor its commitments under a trade agreement, if it perceives that the short-term benefit of "cheating" is less than the long-term cost of the ensuing retaliatory measures that its trading partner would then adopt. We thus characterize self-enforcing trade agreements as equilibria of a repeated or dynamic game with private information.

In the sections that follow, we assume that governments seek to establish a "simple" trade agreement in which they negotiate bound tariff rates. In much of our analysis, we consider

[14] In particular, suppose that governments may make transfers to one another at the same time that they apply their respective tariffs, where transfers are lump-sum instruments that do not affect joint government welfare. Consider the tariff-transfer scheme $\{\tau(\gamma), T(\gamma)\}$ where $\tau(\gamma) = \tau^E(\gamma)$ and $T(L) = 0 < T(H) = W_x(\tau^E(H); L) - W_x(\tau^E(L); L)$. This scheme is incentive compatible and implements efficient tariffs. Intuitively, only a government with a high type will apply the efficient tariff for a high type, if the application of that tariff requires as well that a transfer of sufficient size be made to the other government. Bagwell and Staiger (2005) characterize the efficient tariff-transfer scheme for a model with a continuum of types.

whether governments can enforce a trade agreement in which they apply the tariffs that are induced by the optimal weak binding. Throughout our analysis, in the event that a government cheats and applies a tariff in excess of the bound rate, we assume that governments abandon cooperation and apply Nash tariffs thereafter. Of course, more sophisticated (e.g., carrot-stick) punishment schemes could be entertained as well; however, if governments require an explicit agreement in order to coordinate on a cooperative path, it is perhaps unlikely that a breakdown in cooperation would be consistent with the coordination of a sophisticated punishment path.[15]

In the present section, we begin our formal analysis with a model in which each government's privately observed political-economy parameter (i.e., its type) is independently determined over time. Bagwell and Staiger (2005) study a similar dynamic model when a continuum of possible types exists. Here, we highlight the additional predictions that are available in our simple two-type model. In subsequent sections, we build on this foundation and allow that the privately observed types are perfectly and imperfectly persistent through time.

4.1. The Repeated Game with Transitory Shocks

The *repeated game with transitory shocks* is informally described as follows. In each period, each of the two governments privately observes its type, where types are identically and independently distributed across governments and over time. The governments then simultaneously apply their tariffs, and government welfare for the period is determined. Applied tariffs are publicly observed. Thus, at the start of any period t, each government observes a private history and a public history. A government's private history includes its knowledge of its current and past types. The public history that government's share is their common observation of the tariffs applied by both governments in all preceding periods. The same sequence repeats in each of an infinite number of periods, $t = 1, ..., \infty$.

We consider the *perfect public equilibria* (Fudenberg, Levine and Maskin, 1994) of this game. Hence, we require that each government uses a *public strategy*: the tariff that a government applies in period t may depend on the public history of applied tariffs and on this government's private observation of its current type; however, a government's applied tariff in period t may not depend upon payoff-irrelevant private information, such as its private recollection of its types in previous periods. We also restrict attention to *symmetric perfect public equilibria*, which we define as perfect public equilibria in which governments use ex ante symmetric strategies. Thus,

[15] Following Bagwell and Staiger (2002, Chapter 6; 2005), we regard retaliation provisions in the WTO Dispute Settlement Understanding as facilitating on-equilibrium-path rebalancing of tariffs in response to political-economic shocks. By contrast, the Nash reversion that follows an off-equilibrium-path deviation captures the understanding among governments that non-compliant behavior could lead to a breakdown in the cooperative endeavor. In the two-type model considered here, the scope for rebalancing provisions is limited. See Bagwell and Staiger (2005) for a model with a continuum of types in which WTO escape clause rules can be interpreted as on-equilibrium-path responses to political-economic shocks. Further discussion of the roles of retaliation in the WTO, see also Bagwell (forthcoming).

in a symmetric perfect public equilibrium, strategies are exchangeable across governments as a function of public histories and current types and do not depend on governments' "names." Throughout, we refer to symmetric perfect public equilibria as *equilibria*.

In repeated games with private information, a player may consider two kinds of deviations.[16] A first kind of deviation is known as an *on-schedule deviation*. A deviation of this kind occurs when a government has type $\gamma \in \{L, H\}$ in some period t and deviates by applying the tariff that is intended for type $\gamma' \in \{L, H\}$, where $\gamma' \neq \gamma$ and the equilibrium tariffs for L and H in period t are distinct. An on-schedule deviation by one government is not observable, as a deviation, to the other government. The second kind of deviation is called an *off-schedule deviation*. A deviation of this kind occurs in period t if a government applies a tariff that is not prescribed by the equilibrium for either of the two possible types of governments. An off-schedule deviation by one government is observable, as a deviation, to the other government. In the event of an off-schedule deviation, play has moved off of the equilibrium path, and governments may then move into a punishment phase.

We assume that governments negotiate a trade agreement at date zero, before learning their first-period types. In the present section, we assume further that the agreement takes the form of a weak binding. The associated applied tariffs can be enforced in an equilibrium of the repeated game with transitory shocks if they do not invite on-schedule or off-schedule deviations. We define below a class of "weak-binding equilibria" for the repeated game, in which the applied tariffs induced by a weak binding do not give rise to an on-schedule deviation. For such equilibria, the important issue, instead, is whether governments are sufficiently patient to resist undertaking an off-schedule deviation.

4.2. Weak-Binding Equilibria

We define a *weak-binding equilibrium* as an equilibrium in which there exists a weak binding τ^W such that (i) along the equilibrium path, governments apply tariffs that are equal to or below τ^W, and (ii) at any period t, the applied tariffs selected by governments are independent of the public history so long as no government has previously applied a tariff in excess of τ^W. The first requirement is simply a definition of a weak binding. If a government applies a tariff that exceeds the bound rate, it has violated the agreement. The second requirement goes somewhat further and posits that any applied tariff at or below the bound rate is "fine," in the sense that the applied tariffs selected in the future by one government are not sensitive to the particular way in which the other government has met its weak-binding obligation.

With the notion of a weak-binding equilibrium, we thus have a way to formalize the appealing idea that an agreement to use a weak binding effectively partitions the public history of applied tariffs into "good histories" (both governments have historically applied tariffs that

[16] Our discussion here follows Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004).

satisfy the agreed-upon weak binding) and "bad histories" (at least one government has applied a tariff that violates the weak binding). In the latter event, some government has taken an off-schedule deviation in which it applied a tariff in excess of τ^W, and governments then abandon cooperation and revert to Nash tariffs in all future periods. Notice, though, that governments do not punish (or reward) an off-schedule deviation that involves an applied tariff that is at or below the weak binding. In any period along the equilibrium path, governments can thus choose freely over all applied tariffs that satisfy the weak binding. It follows that, in a weak-binding equilibrium, τ^W induces the applied tariff schedule $\tau(\gamma, \tau^W) = \min\{\tau^W, \tau^N(\gamma)\}$, where our notation now makes explicit the dependence of the applied tariff schedule on the weak binding. As discussed in the previous section, the induced applied tariffs trivially satisfy the incentive-compatibility constraint for the static game. We may thus conclude that the applied tariffs induced by a weak-binding equilibrium do not give rise to an on-schedule deviation.

The more interesting issue is whether governments are willing to forego off-schedule deviations in which they apply tariffs in excess of the weak binding. In the remainder of this section, we characterize the critical discount factor above which governments can support the applied tariffs induced by the optimal weak binding, τ_A^W, in a weak-binding equilibrium of the repeated game with transitory shocks. We also characterize the "most-cooperative" polices that can be enforced in a weak-binding equilibrium when governments are less patient.

To start, we consider any weak binding, τ^W, and characterize the short-term incentive to cheat and the long-term discounted cost of a breakdown in cooperation, respectively. We define the short-term gain from cheating from an applied tariff τ for a government with type γ as

$$\Omega(\tau; \gamma) \equiv W_x(\tau^N(\gamma); \gamma) - W_x(\tau; \gamma). \tag{4.1}$$

Notice that $\Omega(\tau; \gamma)$ is strictly convex in τ and reaches a minimum of zero when $\tau = \tau^N(\gamma)$.

Consider now how the incentive to cheat varies with the weak binding, τ^W, for a government with type γ. If governments use a weak binding, τ^W, the induced applied tariff schedule is $\tau(\gamma, \tau^W) = \min\{\tau^W, \tau^N(\gamma)\}$. Thus, a government with type γ would evaluate the potential benefits of cheating from the applied tariff $\tau(\tau^W, \gamma)$:

$$\Omega(\tau(\gamma, \tau^W); \gamma) \equiv W_x(\tau^N(\gamma); \gamma) - W_x(\tau(\gamma, \tau^W); \gamma). \tag{4.2}$$

We distinguish between two cases. First, if $\tau^W \geq \tau^N(\gamma)$, then $\tau(\gamma, \tau^W) = \tau^N(\gamma)$ and so $\Omega(\tau(\gamma, \tau^W); \gamma) = 0$. A government has no incentive to cheat when the weak binding exceeds its Nash tariff, as it then already applies its optimal tariff. Second, if $\tau^W < \tau^N(\gamma)$, then $\tau(\gamma, \tau^W) = \tau^W$, and an incentive to cheat thus exists. In Figure 6a, we illustrate $\Omega(\tau(\gamma, \tau^W); \gamma)$ as a function of τ^W. A case of special interest occurs when governments agree to set the weak binding at its optimal level, $\tau_A^W = \tau^E(H)$. The induced applied tariffs are then $\tau(L, \tau_A^W) = \tau^N(L)$ and $\tau(H, \tau_A^W) = \tau^E(H)$; hence, we have that $\Omega(\tau(L, \tau_A^W); L) = 0$ and $\Omega(\tau(H, \tau_A^W); H) = W_x(\tau^N(H); H) - W_x(\tau^E(H); H) > 0$. Thus, when governments agree to use the optimal weak binding, an incentive to cheat exists only for a government with the high type.

We consider next the future value of cooperation. If cheating occurs, then in each future period governments abandon cooperation and apply Nash tariffs. For the repeated game with transitory shocks, when a weak binding τ^W is used, the per-period value of cooperation is

$$\omega(\tau^W) \equiv E[W_x(\tau(\gamma,\tau^W);\gamma) - W_x(\tau^N(\gamma);\gamma)] + E[W_x^*(\tau(\gamma,\tau^W)) - W_x^*(\tau^N(\gamma))] \tag{4.3}$$

where the expectation is taken over γ. The first term is negative and corresponds to the fact that a government gains with respect to its import good from a breakdown in cooperation, since it then applies its optimal tariff regardless of its type. The second term captures the potential loss to a government from a breakdown in cooperation: its export good then confronts Nash tariffs abroad.[17] Recalling the definition of joint welfare $J(\tau;\gamma)$ and using (4.3), we observe that the per-period value of cooperation can be re-written as

$$\omega(\tau^W) \equiv E[J(\tau(\gamma,\tau^W);\gamma) - J(\tau^N(\gamma);\gamma)] \tag{4.4}$$

where the expectation is taken over γ.

We consider how $\omega(\tau^W)$ varies with τ^W. For $\tau^W \geq \tau^N(H)$, the weak binding exceeds all Nash tariffs, and so the induced applied tariff schedule is simply the Nash tariff schedule: $\tau(\gamma,\tau^W) = \tau^N(\gamma)$. It follows that $\omega(\tau^W) = 0$ for all $\tau^W \geq \tau^N(H)$. For $\tau^W \in [\tau^N(L), \tau^N(H))$, the induced applied tariff for the low type remains $\tau^N(L)$; however, the induced applied tariff for the high type is now τ^W. As we argue above, the optimal weak binding is thus $\tau_A^W = \tau^E(H)$. Hence, for $\tau^W \in [\tau^N(L), \tau^N(H))$, $\omega(\tau^W)$ is strictly concave and obtains its maximum at $\tau_A^W = \tau^E(H)$, at which point $\omega(\tau^E(H)) > 0$ by Proposition 3. Over this range, $\omega(\tau^W) = \eta_H[J(\tau^W;H) - J(\tau^N(H);H)]$, since the low type applies its Nash tariff whether or not governments cooperate. Thus, the sign of $\omega(\tau^N(L))$ is the same as the sign of $J(\tau^N(L);H) - J(\tau^N(H);H)$. Using (3.4), we thus have that

$$sign\{\omega(\tau^N(L))\} = sign\{L - L^N(H)\}. \tag{4.5}$$

As shown in Figure 3, it follows that $\omega(\tau^N(L)) > 0$ if $H - L$ is sufficiently small. Finally, if $\tau^W < \tau^N(L)$, then the induced applied tariff is $\tau(\gamma,\tau^W) = \tau^W$ for $\gamma = L$ and $\gamma = H$. The weak binding is then equivalent in its effect to a strong binding. Since the optimal strong binding is $\tau_A^S = \tau^E(E\gamma) > \tau^N(L)$ under our parameter restriction, $\omega(\tau^w)$ is strictly increasing in τ^W for $\tau^W < \tau^N(L)$. The function $\omega(\tau^w)$ is illustrated in Figure 6b for the case in which $\omega(\tau^N(L)) > 0$.

We may now state and analyze the off-schedule incentive constraint for a weak-binding equilibrium in the repeated game with transitory shocks. Since types are independent over time, a government's expected future discounted value of cooperation is independent of its current type. We assume that governments discount the future using a common discount

[17] Recall that the symmetric structure of our model enables us to focus on good x, since the expected welfare for the government of the home country on good y is the same as the expected welfare for the government of the foreign country on good x.

factor, $\delta \in (0,1)$. Hence, if governments agree to use a weak binding, τ^W, the off-schedule incentive constraint is as follows:

$$\Omega(\tau(\gamma, \tau^W); \gamma) \leq \frac{\delta}{1-\delta}\omega(\tau^W), \tag{4.6}$$

where $\Omega(\tau(\gamma, \tau^W); \gamma)$ and $\omega(\tau^W)$ are defined in (4.2) and (4.4), respectively. The applied tariffs induced by a weak binding τ^W can be enforced in a weak-binding equilibrium if and only if (4.6) holds.

We now characterize the critical discount factor, δ_A^W, above which governments can support a weak-binding equilibrium in which the weak binding is set at the optimal weak binding, τ_A^W. At this binding, a government with a low type applies its optimal tariff, $\tau^N(L)$, and has no incentive to cheat. A government with a high type, by contrast, has a positive incentive to cheat. Thus, (4.6) holds for the optimal weak binding, τ_A^W, if and only if $\Omega(\tau_A^W; H) \leq \frac{\delta}{1-\delta}\omega(\tau_A^W)$. Given $\tau_A^W = \tau^E(H)$, we find that (4.6) holds for the optimal weak binding if and only if

$$\delta \geq \delta_A^W \equiv \frac{W_x(\tau^N(H), H) - W_x(\tau^E(H), H)}{W_x(\tau^N(H), H) - W_x(\tau^E(H), H) + \eta_H[J(\tau^E(H); H) - J(\tau^N(H); H)]}. \tag{4.7}$$

Using (4.7), it is easy to see that $\delta_A^W \in (0,1)$. Thus, for sufficiently patient governments, an agreement to apply the tariffs induced by the optimal weak binding is self enforcing, where the critical level of patience is given by δ_A^W.

The critical discount factor is a function of model parameters. Using (4.7), we see that δ_A^W is differentiable and strictly decreasing in η_H. We notice as well that δ_A^W approaches unity as η_H approaches zero. Intuitively, if a government has a high type today and believes that it and, most importantly, the other government will almost always have low types in future periods, then the threat of a Nash trade war is essentially meaningless, since the other government is almost sure to apply the tariff $\tau^N(L)$ in future periods even under cooperation. In this situation, a government with a high type can resist the incentive to cheat only if it has almost limitless patience. At the other extreme, when η_H approaches unity, δ_A^W is low but still positive. This reflects the idea that the cost of a Nash trade war is greatest when the other government is expected to almost always have a high type in future periods, since the applied tariff of the other government would then almost always rise in future periods were cooperation abandoned.

The determination of δ_A^W is illustrated in Figure 7. We depict there the off-schedule incentive constraint for the high type of government under different discount factors. As illustrated, when $\delta = \delta_A^W$, a government with the high type is just indifferent between applying a tariff at the bound level τ_A^W and cheating. The discount factor δ_B is big in the sense that $\delta_B > \delta_A^W$. If a government with a high type has this level of patience, the incentive to cheat when the bound tariff is τ_A^W is smaller than the discounted value of cooperation. Finally, the discount factor δ_S is small in the sense that $\delta_S < \delta_A^W$. For a small discount factor, governments are sufficiently impatient that a government with the high type would cheat on an agreement to honor the binding τ_A^W. Before discussing the agreement that less patient governments would strike, we first introduce some definitions.

For a given $\delta \in (0,1)$, we define a weak binding τ^W as an *equilibrium binding* if there exists a weak-binding equilibrium in which governments use the weak binding τ^W. For a given $\delta \in (0,1)$, we next define a weak binding τ^W as a *most-cooperative equilibrium binding* if no other equilibrium binding exists that offers higher expected joint welfare. Formally, the most-cooperative equilibrium binding is the weak binding τ^W that solves the following program:

$$\max_{\tau^W} EJ(\tau(\gamma, \tau^W); \gamma) \text{ subject to (4.6).}$$

Equivalently, the most-cooperative equilibrium binding is the τ^W that maximizes $\omega(\tau^W)$ subject to (4.6). We denote the most-cooperative equilibrium binding as $\tau^W(\delta)$.

As our discussion above suggests, for $\delta \geq \delta_A^W$, the most-cooperative equilibrium binding is simply the optimal weak binding: $\tau^W(\delta) = \tau_A^W$. As Figure 7 illustrates, for $\delta < \delta_A^W$, we may characterize $\tau^W(\delta)$ as the lowest τ such that $\Omega(\tau; H) = \frac{\delta}{1-\delta}\omega(\tau)$.[18] Notice that $\tau^W(\delta) \in (\tau_A^W, \tau^N(H))$ for $\delta \in (0, \delta_A^W)$ and that $\tau^W(\delta)$ approaches $\tau^N(H)$ as δ goes to zero. It is straightforward to verify that $\tau^W(\delta)$ is differentiable and strictly decreasing in $\delta \in (0, \delta_A^W)$. The applied tariffs induced by the most-cooperative equilibrium binding are given by the schedule $\tau(\gamma, \tau^W(\delta))$. These applied tariffs are illustrated in Figure 8 as a function of the discount factor. Notice that as governments become less patient, the most-cooperative equilibrium binding rises, and this enables a government with a high type to apply a higher tariff. In turn, when this government applies a higher tariff, its incentive to cheat is diminished. In this general way, less patient governments can achieve some cooperation by raising the weak binding just enough to dissuade a high-type government from cheating.

We may now state the following proposition:

Proposition 7: *Consider the repeated game with transitory shocks. (a) There exists $\delta_A^W \in (0,1)$ such that, for all $\delta \geq \delta_A^W$, the most-cooperative equilibrium binding is the optimal weak binding: $\tau^W(\delta) = \tau_A^W$. (b) For $\delta \in (0, \delta_A^W)$, the most-cooperative equilibrium binding is $\tau^W(\delta) \in (\tau_A^W, \tau^N(H))$, where $\tau^W(\delta)$ is differentiable and strictly decreasing in δ. (c) The critical discount factor, δ_A^W, is differentiable and strictly decreasing in η_H and approaches unity as η_H approaches zero.*

Proposition 7 summarizes our formal analysis of the repeated game with transitory shocks. Bagwell and Staiger (2005) find that part (a) also holds in a model with a continuum of types; however, as they observe, part (b) may fail in their model, since $\tau^W(\delta)$ may be discontinuous. Finally, part (c) reflects our two-type formulation and delivers the interesting implication

[18] For $\delta \in (0,1)$, this equation has two roots, and the higher root is $\tau = \tau^N(H)$. We are interested in the lower root, since it is closer to the most efficient tariff for a government with the highest type and thus results in greater expected joint welfare.

that cooperation is easier (i.e., δ_A^W is lower) among governments that often face *high* political-economy pressures. We will see that related themes arise in subsequent sections, when we allow for persistent types.

5. Enforcement: Perfectly Persistent Shocks

In the preceding section, we assume that governments privately observe political-economic pressures that are transitory in nature. In this section, we maintain the assumption that each government has some private information about its political-economic pressures; however, we now relax the assumption that these pressures are transitory. While the assumption of transitory shocks simplifies the analysis, it seems more realistic to allow that political-economic pressures exhibit some persistence. Once persistence is introduced, the analysis becomes more advanced, since a government's tariff in the current period may reveal its current type and thereby reveal information about the government's probable future type. In the current section, we relax the assumption of transitory shocks by taking the opposite extreme case and supposing that a government's type is perfectly persistent through time. In the next section, we discuss the case of imperfectly persistent shocks.

5.1. The Dynamic Game with Perfectly Persistent Shocks

The *dynamic game with perfectly persistent shocks* is informally described as follows. At the start of the game, each government privately observes its type, where types are independently distributed across governments and perfectly persistent (i.e., fixed) through time. The game is otherwise the same as the repeated game with transitory shocks. Thus, in any period t, governments simultaneously apply their tariffs, given the public history of previous applied tariffs and governments' private observations as to their respective types. Since types are now persistent over time, the game is no longer a repeated game. It is instead a dynamic game.

As our solution concept, we consider *perfect bayesian equilibria* (Fudenberg and Tirole, 1991) that satisfy the further refinement that each government uses a public strategy. In turn, a government uses a public strategy if any period its applied tariff depends only on the public history at that time and the its payoff-relevant private information. For our model, a government's payoff-relevant private information is its type, $\gamma \in \{L, H\}$. Following Athey and Bagwell (2006), we may refer to such refined equilibria as *perfect public bayesian equilibria*. Perfect public bayesian equilibrium is a natural extension of perfect public equilibrium to dynamic bayesian games as opposed to repeated games. In the dynamic game with persistent types, perfect public bayesian equilibria are defined in terms of the strategy that each government uses in the dynamic game and the belief function that each government uses when making inferences about the other government's type. In a perfect public bayesian equilibrium, each government's public strategy must be sequentially rational and its belief function must be bayesian where possible given the public history of observed tariffs and its understanding of the other government's equilibrium strategy.

Since governments are ex ante symmetric, we also restrict attention to *symmetric perfect public bayesian equilibria*. As in the previous section, the requirement of symmetry means that strategies are exchangeable across governments and do not depend on governments' "names." Even though governments use ex ante symmetric strategies, asymmetries in governments' strategies may emerge over time, as a result of asymmetric type realizations and past behavior. We refer to symmetric perfect public equilibria as *equilibria*.

In the dynamic game with perfectly persistent types, we may again distinguish between on-schedule and off-schedule deviations. Following Athey and Bagwell (2006), an on-schedule deviation is now a deviation in which a government with type $\gamma \in \{L, H\}$ mimics the behavior that it would take were its type $\gamma' \in \{L, H\}$, where $\gamma \neq \gamma'$, in all periods of the game and for all public histories. An off-schedule deviation occurs whenever a government applies tariffs that are not consistent with the equilibrium strategies of either type of government.

We assume that governments negotiate a trade agreement at date zero, before they learn their respective types for the dynamic game. The probability that a government receives a high (low) type for the dynamic game is η_H (η_L). We focus again on the possibility of an agreement among governments to use a weak binding. In previous sections, we define the applied tariffs that are induced by an agreement to use a weak binding. Here, we consider whether these induced applied tariffs can be enforced in an equilibrium of the dynamic game with perfectly persistent types. We identify a problem that may be associated with the enforcement of these tariffs, when types are perfectly persistent. We then consider alternative applied tariffs that can be enforced when a weak binding is used.

5.2. Weak-Binding Equilibria

We again assume that governments negotiate an agreement to use a weak binding. As in the previous section, we define a *weak-binding equilibrium* as an equilibrium in which there exists a weak binding, τ^W, such that (i) along the equilibrium path, governments apply tariffs that are equal to or below τ^W, and (ii) at any period t, the applied tariffs selected by governments are independent of the public history so long as no government has previously applied a tariff in excess of τ^W. Recall that the second requirement posits that any applied tariff at or below the weak binding is "fine." Thus, in a weak-binding equilibrium, the applied tariff schedule $\tau(\gamma, \tau^W) = \min\{\tau^W, \tau^N(\gamma)\}$ is induced in all periods along the equilibrium path.

In a weak-binding equilibrium, if a government undertakes an off-schedule deviation in which it applies a tariff in excess of τ^W, then governments abdandon cooperation and revert to Nash tariffs in all future periods. Importantly, in the dynamic game with persistent types, a government's expected welfare under Nash play depends upon the government's type and the government's belief about the other government's type. If one government believes that the other government probably has a high (low) type, then the former government expects that a Nash punishment would probably mean that its exporters would face the tariff $\tau^N(H)$ ($\tau^N(L)$). A government's belief must be bayesian in equilibrium, and so a government's expected value of future cooperation may be influenced by the particular applied tariffs selected by the other government. Thus, while any applied tariff below the weak binding is fine, it is unavoidable that one government's particular selections may affect the other government's bayesian beliefs about the former government's type.

We distinguish between two ranges for the weak binding. Suppose first that $\tau^W > \tau^N(L)$. This range includes the optimal weak binding, τ_A^W. When $\tau^W > \tau^N(L)$, a weak-binding equilibrium is possible only if a government with a low type applies the tariff $\tau(L, \tau^W) = \tau^N(L)$ in all periods along the equilibrium path. Likewise, if $\tau^W \geq \tau^N(H)$, then a government with a high type applies the tariff $\tau(H, \tau^W) = \tau^N(H)$ in all periods along the equilibrium path. We conclude that there exists a weak-binding equilibrium in which $\tau^W \geq \tau^N(H)$, and in any such equilibrium a government of type $\gamma \in \{L, H\}$ applies its Nash tariff $\tau^N(\gamma)$ in every period. Thus, a weak-binding equilibrium with $\tau^W \geq \tau^N(H)$ exists but does not improve on the no-agreement Nash equilibrium benchmark. If instead $\tau^W \in (\tau^N(L), \tau^N(H))$, then a weak-binding equilibrium exists only if a government with a high type applies the tariff $\tau(H, \tau^W) = \tau^W$ in all periods along the equilibrium path. Hence, if $\tau^W \in (\tau^N(L), \tau^N(H))$, then a weak-binding equilibrium can exist only if, in all periods along the equilibrium path, a government with a low type applies the tariff $\tau^N(L)$ and a government with a high type applies the tariff τ^W.

Let us now consider the off-schedule incentive constraint that arises for a government with a high type in a weak-binding equilibrium when $\tau^W \in (\tau^N(L), \tau^N(H))$. In particular, suppose that one government has a high type while the other government has a low type. In period one, the former government necessarily applies the tariff τ^W while the latter government necessarily applies the tariff $\tau^N(L)$. At this point, the governments can infer each other's types, and the former government thus knows that the latter government has a low type. In the second period, a government with a high type then cannot resist taking an off-schedule deviation in which it violates the binding and applies its Nash tariff, $\tau^N(H)$. With respect to its import good, this government gains from such an off-schedule deviation, since it thereby applies its Nash tariff in the current and all future periods. With respect to its export good, this government does not gain or lose, since it knows that the low-type government applies the tariff $\tau^N(L)$ in period two and all future periods both on and off of the equilibrium path (i.e., whether the governments continue cooperating or revert to an infinite Nash punishment). Hence, it cannot be true that a government with a high type applies the tariff $\tau^W \in (\tau^N(L), \tau^N(H))$ in all periods along the equilibrium path, and we therefore conclude that a weak-binding equilibrium fails to exist when $\tau^W \in (\tau^N(L), \tau^N(H))$.

The following proposition summarizes our discussion to this point:

Proposition 8: *Consider the dynamic game with perfectly persistent shocks. (i). There exists* *a weak-binding equilibrium with* $\tau^W \geq \tau^N(H)$. *In any such equilibrium, a government of type* $\gamma \in \{L, H\}$ *applies its Nash tariff* $\tau^N(\gamma)$ *in every period. (ii). There does not exist a weak-binding equilibrium with* $\tau^W \in (\tau^N(L), \tau^N(H))$.

Part (ii) of this proposition is most interesting and captures a "no-good-deed-goes-unpunished" phenomenon. When a government has low political-economic pressures and thus applies a tariff below the bound level, the other government gains from this good deed. At the same time, the other government then realizes that its partner has a low type and thus cannot credibly threaten to raise its tariffs in response to a violation. Given this realization, the other government cannot resist "punishing" the good deed by cheating and applying a tariff in excess of the weak binding. In effect, when types are perfectly persistent, a "ratchet effect" exists with respect to the enforcement of an agreement to use a weak binding: once a government reveals that it has the weak type, the other government cannot resist undertaking actions that hurt the former government.[19]

It is interesting to compare Propositions 7 and 8. When types are transitory, a government with a low type today may have a high type tomorrow. Thus, when types are transitory and $\tau^W \in (\tau^N(L), \tau^N(H))$, a government that cheats today is sure to suffer some retaliatory cost in the future, since the other government will apply $\tau^N(H)$ rather than τ^W in all future periods in which it has a high type. This explains why, as Proposition 7 confirms, sufficiently patient governments that face transitory shocks can enforce an agreement to abide by the optimal weak binding. By contrast, as Proposition 8 establishes, when types are perfectly persistent, a government that reveals itself to be "weak" in the sense of having low political-economic pressures reveals as well that it cannot impose a retaliatory cost on a cheating government.

We turn now to the second possible range for the weak binding; specifically, we suppose that governments agree to set a weak binding that is equal to or below the Nash tariff for the low type: $\tau^W \leq \tau^N(L)$. Given that the weak binding falls at or below the Nash tariffs for both the low and high types of government, a weak-binding equilibrium can exist only if both types of government apply the tariff τ^W in all periods along the equilibrium path.

To examine the off-schedule incentive constraint, we begin with the short-term incentive to cheat. The most attractive off-schedule deviation for a government with type γ is its Nash tariff, $\tau^N(\gamma)$. Using (4.1), the short-term incentive to cheat for a government with type γ is thus $\Omega(\tau^W; \gamma)$. It is direct to confirm that the incentive to cheat is greatest for a government with the high type: $\Omega(\tau^W; H) - \Omega(\tau^W; L) = [W_x(\tau^N(H); H) - W_x(\tau^W; H)] - [W_x(\tau^N(L); L) - W_x(\tau^W; L)] > [W_x(\tau^N(L); H) - W_x(\tau^W; H)] - [W_x(\tau^N(L); L) - W_x(\tau^W; L)] \geq 0$, where the final inequality follows since $f' > 0$ and $\tau^N(L) \geq \tau^W$.

Consider now the future value of cooperation. If a government cheats in period t, then in all future periods governments apply their Nash tariffs rather than the bound tariff, τ^W. In

[19] The ratchet effect arises in dynamic contracting problems with incomplete information in which long-term contracts are infeasible. See, for example, Freixas, Guesnerie and Tirole (1985) and Laffont and Tirole (1988).

contrast to our characterization in (4.3) of the per-period value of cooperation in the repeated game with transitory shocks, we must now allow that the per-period value of cooperation depends upon a government's persistent type. In our dynamic game with perfectly persistent shocks, when governments seek to enforce a weak-binding equilibrium and $\tau^W \leq \tau^N(L)$, we thus define the per-period value of cooperation for a government with type γ as

$$\widetilde{\omega}(\tau^W;\gamma) = [W_x(\tau^W;\gamma) - W_x(\tau^N(\gamma);\gamma)] + E[W_x^*(\tau^W) - W_x^*(\tau^N(\gamma))], \tag{5.1}$$

where the expectation is taken over γ. Thus, when evaluating the per-period value of cooperation, a government considers its own persistent type and recognizes that with probability η_H (η_L) the other government will select the tariff $\tau^N(H)$ ($\tau^N(L)$) if cooperation breaks down. The first term in (5.1) is negative (and strictly so if $\tau^W < \tau^N(\gamma)$), due to the benefit that a government enjoys on its import good when applying its optimal tariff. The second term in (5.1) is positive when $\tau^W \leq \tau^N(L)$, since a government then experiences a welfare loss on its export good when its trading partner abandons cooperation and applies Nash tariffs.

Rearranging terms slightly, we now confirm that the per-period value of cooperation is lower for a government with a high type: $\widetilde{\omega}(\tau^W;H) - \widetilde{\omega}(\tau^W;L) = \Omega(\tau^W;L) - \Omega(\tau^W;H) < 0$, where the inequality follows as above from $f' > 0$ and $\tau^N(L) \geq \tau^W$. Intuitively, whether a government has a low or high type, it faces the same welfare loss when the other government abandons cooperation and imposes its Nash tariffs. The reason that the per-period value of cooperation varies with a government's type is that a high type of government gains more from abandoning τ^W and applying its Nash tariff when $\tau^W \leq \tau^N(L)$.

For our dynamic game with perfectly persistent shocks, if governments seek to enforce a weak-binding equilibrium and $\tau^W \leq \tau^N(L)$, then the off-schedule incentive constraint is

$$\forall \gamma \in \{L, H\},\ \Omega(\tau^W;\gamma) \leq \frac{\delta}{1-\delta}\widetilde{\omega}(\tau^W;\gamma). \tag{5.2}$$

Given $\tau^W \leq \tau^N(L)$, we confirm above that the LHS of this constraint is largest when $\gamma = H$ and that the RHS is smallest when $\gamma = H$. Thus, when $\tau^W \leq \tau^N(L)$, (5.2) holds if and only if the off-schedule constraint holds for a government with the high type:

$$\Omega(\tau^W;H) \leq \frac{\delta}{1-\delta}\widetilde{\omega}(\tau^W;H). \tag{5.3}$$

With the off-schedule incentive constraint thus represented, we are in a position to characterize the conditions under which there exist weak-binding equilibria with $\tau^W \leq \tau^N(L)$.

To this end, we rearrange (5.1) and observe that the per-period value of cooperation for a government with a high type can be written as

$$\widetilde{\omega}(\tau^W;H) = [J(\tau^W;H) - J(\tau^N(H);H)] + \eta_L[W_x^*(\tau^N(H)) - W_x^*(\tau^N(L))]. \tag{5.4}$$

When $\tau^W \leq \tau^N(L)$, our parameter restriction yields $\tau^W \leq \tau^N(L) < \tau^E(H)$; thus, when $\tau^W \leq \tau^N(L)$, we have that $J(\tau^W;H) \leq J(\tau^N(L);H)$. Given that the second term in (5.4) is clearly negative, the following conclusion is now apparent: when $\tau^W \leq \tau^N(L)$, if $J(\tau^N(L);H) \leq J(\tau^N(H);H)$, then $\widetilde{\omega}(\tau^W;H) < 0$. Hence, if $J(\tau^N(L);H) \leq J(\tau^N(H);H)$, then the off-schedule incentive constraint (5.3) cannot hold when $\tau^W \leq \tau^N(L)$, and so in this case there does not exist a weak-binding equilibrium with $\tau^W \leq \tau^N(L)$.

The other possibility is that $\tau^W \leq \tau^N(L)$ and $J(\tau^N(L);H) > J(\tau^N(H);H)$. In this event, we may conclude from (5.4) that $\widetilde{\omega}(\tau^W;H) > 0$ if τ^W is sufficiently close to $\tau^N(L)$ and η_L is sufficiently small. Let us set $\tau^W = \tau^N(L)$ and therby maximize the opportunity to achieve a positive value for $\widetilde{\omega}(\tau^W;H)$ when $\tau^W \leq \tau^N(L)$. Using (5.4) and rearranging, we find that $\widetilde{\omega}(\tau^N(L);H) > 0$ if and only if

$$\eta_H > \underline{\eta}_H \equiv 1 - \frac{J(\tau^N(L);H) - J(\tau^N(H);H)}{W_x^*(\tau^N(L)) - W_x^*(\tau^N(H))} \tag{5.5}$$

Referring to (5.5), we see that $\underline{\eta}_H \in (0,1)$ if and only if $J(\tau^N(L);H) > J(\tau^N(H);H)$. We conclude that, if $\eta_H > \underline{\eta}_H$, then there exists some critical discount factor $\underline{\delta} \in (0,1)$ such that, for all $\delta \in (\underline{\delta},1)$, (5.3) holds when $\tau^W = \tau^N(L)$.

Building slightly from the preceding discussion, we can now state the following proposition:

Proposition 9: *Consider the dynamic game with perfectly persistent shocks. (i) If $J(\tau^N(L);H) \leq J(\tau^N(H);H)$, or if $J(\tau^N(L);H) > J(\tau^N(H);H)$ and $\eta_H \leq \underline{\eta}_H$, then there does not exist a weak-binding equilibrium when $\tau^W \leq \tau^N(L)$. (ii) If $J(\tau^N(L);H) > J(\tau^N(H);H)$ and $\eta_H \in (\underline{\eta}_H,1)$, then there exists $\underline{\delta} \in (0,1)$ such that, for all $\delta \in (\underline{\delta},1)$, a weak-binding equilibrium exists when $\tau^W = \tau^N(L)$.*

Finally, we recall from (3.4) that $J(\tau^N(L);H) - J(\tau^N(H);H)$ takes the same sign as $L - L^N(H)$. Thus, a weak-binding equilibrium exists when $\tau^W \leq \tau^N(L)$ only if $H - L$ is sufficiently small. A weak-binding equilibrium also exists when $\tau^W \leq \tau^N(L)$ only if the probability η_H that a government has a high type is sufficiently large. Intuitively, the threat of retaliation is more effective when it is more likely that a retaliating government would apply a high Nash tariff.

Propositions 8 and 9 identify only a limited sense in which governments may cooperate in a weak-binding equilibrium. As Proposition 8 indicates, governments may set the weak binding above the Nash tariff of a government with a high type, but in this case the agreement simply induces Nash tariffs. Alternatively, as Proposition 9 confirms, for some parameter regions, governments can enforce a weak-binding equilibrium in which they set the weak binding at or below the Nash tariff of a government with a low type. In this case, however, the applied tariff is always set at the bound rate, and the applied tariff is strictly below the level that is efficient when a single tariff is applied (i.e., the applied tariff is below $\tau^E(E\gamma)$).

In the remainder of the paper, we consider ways in which this limitation might be overcome. First, when a weak binding is used that exceeds the Nash tariff for a government with a low type, such a government might recognize that, when it applies its Nash tariff and reveals its type, it reveals that it is "weak" and thus encourages the other government to cheat. This motivates consideration of a different class of equilibria, in which all applied tariffs at or below the weak binding do not necessarily generate the same continuation path. In this class, we focus in the next subsection on pooling equilibria, wherein the low- and high-types of governments apply the same tariff at all dates along the equilibrium path.[20] Second, recall that we examine in this section an extreme case in which types are perfectly persistent. In Section 6, we consider weak-binding equilibria in a more realistic model in which types are imperfectly persistent. Third, in Section 7, we return to the dynamic game with perfectly persistent shocks and consider a class of equilibria in which the weak binding descends gradually over time as a function of the types revealed by applied tariffs in early periods.

5.3. Pooling Equilibria

We now assume that governments negotiate an agreement to use a weak binding, τ^W; however, we relax the second restriction imposed in our definition of a weak-binding equilibrium and now allow that the continuation of play may depend on the particular way in which governments apply tariffs to satisfy the weak-binding obligation. In particular, one theme of our analysis above is that a government may reveal itself to be weak and incapable of effective retaliation, when it applies a tariff strictly below the weak binding. Rather than look weak and invite an off-schedule deviation by its trading partner, a government with a low type might choose instead to apply the bound tariff, even when its optimal tariff is lower. When a government applies the bound tariff whether its type is high or low, its type cannot be inferred; consequently, the other government may hesitate to cheat with an off-schedule deviation, since it is possible that the retaliatory tariff would be the Nash tariff of a high-type government.

Reflecting this line of thought, we thus now analyze *pooling equilibria*, which are equilibria in which in every period along the equilibrium path each government of each type applies the same tariff. We can understand pooling equilibria as being induced by a weak binding, τ^W, if we assume that each government of each type applies the bound tariff exactly, in order not to be perceived as being weak. This interpretation is motivated in the preceding paragraph. Of course, pooling equilibria can also be understood in terms of an agreement to use a strong binding. Emphasizing the first interpretation, we refer to a pooling equilibrium in which governments apply the bound tariff τ^W in all periods along the equilibrium path as a *pooling equilibrium at the weak binding.*

Consider now the off-schedule incentive constraint for a pooling equilibrium at the weak binding, when governments agree to use a weak binding, τ^W, that satisfies $\tau^W \in (\tau^N(L), \tau^N(H))$. Each government understands that if it applies a tariff below the weak binding, then the other government will believe that it is weak (i.e., a low type). In turn, this belief will encourage the

[20]Pooling often arises as well in dynamic contracting problems in which a ratchet effect is present. See, for example, Freixas, Guesnerie and Tirole (1985) and Laffont and Tirole (1988).

other government to cheat in the following period and induce Nash tariffs for the remainder of time. Consequently, if a government applies a tariff below the weak binding in period t, it should go ahead and apply its Nash tariff in period $t+1$ and thereafter. Likewise, each government understands that if it applies a tariff in excess of the weak binding, then its action will be interpreted as cheating, with the result being that both governments apply Nash tariffs in the next and all subsequent periods. Given this construction, the key off-schedule constraint is that, whatever its type, a government must earn greater expected discounted welfare when both governments cooperate and apply the bound tariff, τ^W, throughout time than when the government cheats with an off-schedule deviation to its Nash tariff and thereby induces both governments to apply Nash tariffs in all future periods. Notice that a government with a high type thus contemplates a deviation to a tariff that exceeds the bound rate ($\tau^N(H) > \tau^W$) and is in this sense a violation whereas a government with a low type evaluates a deviation to a tariff that falls below the bound rate ($\tau^N(L) < \tau^W$) and thereby signals weakness.

Formally, for a pooling equilibrium at the weak binding, the off-schedule constraint must hold for both types of government and is given by (5.2). Above, we argue that this constraint binds first for a high type of government, when $\tau^W \leq \tau^N(L)$. We assume in this subsection, however, that $\tau^W \in (\tau^N(L), \tau^N(H))$. Hence, it is not clear whether the off-schedule incentive constraint binds first for a government with a high or low type. If governments anticipate a pooling equilibrium at the weak binding and seek to maximize expected joint welfare, then the most attractive weak binding is $\tau^W = \tau^S_A = \tau^E(E\gamma)$. As in our analysis of the optimal strong binding, when governments anticipate that a single applied tariff will be applied in all periods along the equilibrium path, they prefer that the applied tariff be the tariff that is efficient for the expected type. We thus focus on the off-schedule constraint when $\tau^W = \tau^E(E\gamma)$.

Let us now examine the off-schedule constraint in more detail. For a government of type $\gamma \in \{L, H\}$, the constraint can be written and simplified as follows:

$$\begin{aligned}
\Omega(\tau^W;\gamma) &\leq \frac{\delta}{1-\delta}[W_x(\tau^W;\gamma) + W^*_x(\tau^W) - W_x(\tau^N(\gamma);\gamma) - EW^*_x(\tau^N(\gamma))] \\
&= \frac{\delta}{1-\delta}[-\Omega(\tau^W;\gamma) + W^*_x(\tau^W) - EW^*_x(\tau^N(\gamma))].
\end{aligned}$$

Thus, the off-schedule constraint holds when $\tau^W = \tau^E(E\gamma)$ if and only if

$$\forall\gamma \in \{L, H\},\ \Omega(\tau^E(E\gamma);\gamma) \leq \delta[W^*_x(\tau^E(E\gamma)) - EW^*_x(\tau^N(\gamma))]. \tag{5.6}$$

As established in Proposition 1, the RHS of (5.6) is strictly positive. Notice that the RHS is independent of the deviating government's type and captures the expected welfare loss that the deviating government experiences when the other government applies its Nash tariff rather than the bound tariff. Our next step is to understand how the LHS of (5.6) varies with γ.

Consider then how the incentive to cheat varies with the government's type when $\tau^W = \tau^E(E\gamma)$. At one extreme, if L is only slightly below $L(H, \eta_H)$, then $\tau^E(E\gamma)$ is only slightly above $\tau^N(L)$; thus, the incentive to cheat for a government with the low type is approximately zero. In this case, a government with a high type has the greatest incentive to cheat. At the other extreme, if η_H is close to unitary so that $\tau^E(E\gamma)$ is close to $\tau^E(H)$, and if H is large and close to 7/4 so that $\tau^E(H)$ is close to $\tau^N(H)$, then the incentive to cheat for a government with the high type is approximately zero. In this case, the incentive to cheat is greatest for a government with the low type Thus, when $\tau^W = \tau^E(E\gamma)$, the type of government that has the greatest incentive to cheat varies with the parameters of the model.

To further explore the incentive to cheat, we may pick any (L, H) that lies below the function $L(H, 1)$. We then find $\widetilde{\eta}_H \in (0,1)$ such that $L = L(H, \widetilde{\eta}_H)$. When $\eta_H = \widetilde{\eta}_H$, we have that $\tau^N(L) = \tau^E(E\gamma)$, and our parameter restriction requires that $\eta_H \in (\widetilde{\eta}_H, 1)$. Now, let us define $\Delta_\Omega(\eta_H) \equiv \Omega(\tau^E(E\gamma); H) - \Omega(\tau^E(E\gamma); L)$. We observe first that $\Delta_\Omega(\widetilde{\eta}_H) > 0$, since as noted a government with the low type has zero incentive to cheat when $\tau^N(L) = \tau^E(E\gamma)$. As we increase η_H from $\widetilde{\eta}_H$ toward unity, $\tau^E(E\gamma)$ rises and so it may be expected that $\Delta_\Omega(\eta_H)$ falls. This is in fact the case: straightforward calculations confirm that $\Delta'_\Omega(\eta_H) < 0$. As η_H approaches unity, $\Delta_\Omega(\eta_H) < 0$ is possible; indeed, as noted, if H is close to 7/4, then $\Delta_\Omega(1) < 0$. In this case, the incentive to cheat is greater for a government with a high type if and only if η_H is below some threshold value. For other values of H and L, however, $\Delta_\Omega(1) > 0$. In this case, the incentive to cheat is greater for a government with a high type, for all permissible values of η_H. We can show that $\Delta_\Omega(1) > 0$ if $H - L$ is small in the following sense: $7/4 - H \geq H - L$.

Let us now suppose that the model parameters are such that $\Delta_\Omega(1) > 0$. For all permissible η_H, we then know that the off-schedule constraint is most difficult to satisfy when the government has a high type. Thus, (5.6) holds if and only if the following inequality holds: $\Omega(\tau^E(E\gamma); H) \leq \delta[W_x^*(\tau^E(E\gamma)) - EW_x^*(\tau^N(\gamma))]$. Now, in the limiting case where $\eta_H = 1$, we have that $\tau^E(E\gamma) = \tau^E(H)$ and $EW_x^*(\tau^N(\gamma)) = W_x^*(\tau^N(H))$. In this limiting case, if in addition we put $\delta = 1$, then we can thus rewrite the preceding inequality as $J(\tau^N(H); H) \leq J(\tau^E(H); H)$. Since $\tau^E(H)$ uniquely maximizes $J(\tau; H)$, we conclude that (5.6) holds when $\Delta_\Omega(1) > 0$ if δ and η_H are each sufficiently close to unity.

We summarize our findings in the following proposition:

Proposition 10: *Consider the dynamic game with perfectly persistent shocks. Suppose $\Delta_\Omega(1) > 0$. Then there exists $\underline{\delta} \in (0,1)$ and $\underline{\eta}_H \in (0,1)$ such that for all $\delta \in (\underline{\delta}, 1)$ and for all $\eta_H \in (\underline{\eta}_H, 1)$, there exists a pooling equilibrium at the weak binding, in which the weak binding is $\tau^E(E\gamma)$.*

This proposition illustrates one parameter region over which governments can support a pooling equilibrium in which they always apply the tariff that is efficient for the expected type.[21] As

[21] Recall that $\Delta_\Omega(1) > 0$ holds if $7/4 - H \geq H - L$. Equivalently, $\Delta_\Omega(1) > 0$ holds if $L \geq L(H, 2/3)$. Recall as well that our parameter restriction can be stated as $L(H, \eta_H) > L$. Thus, for any $\eta_H \in (2/3, 1)$, there exist values for H and L such that $\Delta_\Omega(1) > 0$ and our parameter restriction both hold.

our discussion illlustrates, the construction is not trivial, since both the low and high types of government have incentives to cheat.

6. Enforcement: Imperfectly Persistent Shocks

In the preceding two sections, we analyze the two extreme cases in which governments privately observe political-economic pressures that are either transitory or perfectly persistent. Building on the insights obtained in this analysis, we now consider a more realistic model in which governments' types are imperfectly persistent through time. As in the case of perfect persistence, when a government's type exhibits imperfect persistence, its tariff in the current period may reveal its current type and thereby information about its probable future type. We again focus on the applied tariffs that are induced when an agreement is reached to use a weak binding.

6.1. The Dynamic Game with Imperfectly Peristent Shocks

The *dynamic game with imperfectly persistent shocks* is informally described as follows. At the start of the game, each government privately observes its current type, where types are independently distributed across governments and imperfectly persistent through time. As described below, governments' types evolve independently over time according to a Markov process. The game is otherwise the same as those analyzed in the previous two sections. Thus, in any given period, governments simultaneously apply their tariffs, each government has a private history which includes its knowledge of its current type, and goverments share the public history of previous applied tariffs. Since governments' types exhibit some persistence over time, the game is a dynamic (rather than a repeated) game.

A government's type evolves over time according to a simple Markov process. For each government, if the government has the low (high) type in period t, then it has the high (low) type in period $t+1$ with probability λ. We assume that $\lambda \in (0, 1/2)$, so that types exhibit positive persistence through time. Notice that the repeated game with transitory shock corresponds to a setting in which $\lambda = 1/2$ while the dynamic game with perfectly persistent types corresponds to a situation in which $\lambda = 0$.[22] Since types evolve according to a Markov process, a government's knowledge of its own current type (belief as to the other government's current type) suffices for forming an expectation as to its (the other government's) type in the next period.

For our solution concept, we again consider perfect bayesian equilibria (Fudenberg and Tirole, 1991) in which governments use public strategies. Recall that a government uses a public strategy if in any period its applied tariff depends only on the public history at that time and its payoff-relevant private information. Given our Markov assumption, a government's payoff-relevant private information at date t is its current type. We again follow Athey and

[22]When $\eta_H = 1/2$, the repeated game with transitory shocks generates the same distribution of types as does the repeated game with imperfectly persistent shocks when $\lambda = 1/2$.

Bagwell (2006) and refer to such refined equilibria as perfect public bayesian equilibria. In the dynamic game with imperfectly persistent types, perfect public bayesian equilibria are defined in terms of the strategy that each government uses in the dynamic game and the belief function that each government uses when making inferences about the other government's type. In a perfect public bayesian equilibrium, each government's public strategy must be sequentially rational, and its belief function must be bayesian where possible given the public history of observed tariffs and its understanding of the other government's equilibrium strategy and the underlying Markov process that governs the evolution of types. As before, we also assume that governments use ex ante symmetric strategies and thus restrict attention to symmetric perfect public bayesian equilibria. In this section, we refer to symmetric perfect public equilibria for the dynamic game with imperfectly persistent types as *equilibria*.

In the dynamic game with imperfectly persistent types, if in period t a government with one type mimics the behavior that is prescribed in equilibrium for the government when it has the other type, then the government has undertaken an on-schedule deviation. When types are imperfectly persistent, it is possible for a government to undertake an on-schedule deviation in one period and then return to its prescribed strategy in future periods. An off-schedule deviation occurs in the dynamic game with imperfectly persistent types whenever a government applies a tariff in some period t that is not prescribed under the equilibrium strategy for either type that the government may have observed in period t. Even when a government sees that the other governments has undertaken an off-schedule deviation, the former government's beliefs as to evolution over time of the deviating government's type must respect the Markov structure described above.[23]

Once again, we assume that governments negotiate a trade agreement at date zero, before they learn their period-one types for the dynamic game. For simplicity, we assume that a government receives the high type in period one with probability $\eta_H = 1/2$.[24] We focus again on the possibility of an agreement among governments to use a weak binding. In contrast to our findings for the dynamic game with perfectly persistent types, we show that sufficiently patient governments can always enforce an equilibrium in the dynamic game with imperfectly persistent types in which they apply the tariffs, $\tau = \tau(\gamma, \tau^W)$, that are induced by an agreement to use a weak binding, τ^W. In this way, we suggest that concerns about signaling "weakness" will not undermine governments' willingness to apply tariffs below the weak binding, if governments are sufficiently patient, where the critical level of patience is higher when types are more persistent.

[23] The discussion here is informal. See Athey and Bagwell (2006) for formal definitions of equilibrium, on- and off-schedule deviations, and the evolution of beliefs in a related dynamic game with imperfectly persistent types.

[24] This is the stationary distribution for our Markov process. Our parameter restriction can then be specified with respect to this value for η_H.

6.2. Weak-Binding Equilibria

We use the same definition of a weak-binding equilibrium as before. Thus, a weak-binding equilibrium is an equilibrium in which there exists a weak binding, τ^W, such that (i) along the equilibrium path, governments apply tariffs that are equal to or below τ^W, and (ii) at any period t, the applied tariffs selected by governments are independent of the public history so long as no government has previously applied a tariff in excess of τ^W. As argued previously, in a weak-binding equilibrium governments do not punish an off-schedule deviation in which the applied tariff is at or below the weak binding; hence, in a weak-binding equilibrium, the weak binding τ^W induces the applied tariff schedule $\tau(\gamma, \tau^W) = \min\{\tau^W, \tau^N(\gamma)\}$ in all periods along the equilibrium path.

As formalized in Propositions 8 and 9, for the dynamic game with perfectly persistent shocks, weak-binding equilibria often fail to exist; in particular, they do not exist when the weak binding is set at the optimal weak binding, τ_A^W, where $\tau_A^W = \tau^E(H) \in (\tau^N(L), \tau^N(H))$ under our parameter restriction. In that model, a government with a low type that goes below the bound rate and applies its Nash tariff reveals its permanent type; consequently, it reveals as well that it would apply the same tariff in future periods whether or not the other government goes on to cheat and apply a tariff in excess of the binding. We expect this problem to be mitigated somewhat in the dynamic game with imperfectly persistent shocks, since a government that has a low type in one period may have a high type in later periods. Thus, even a government with a low type yesterday can be expected to apply a higher tariff in some future periods, if a violation were to occur today and result in a breakdown in future cooperation.

To explore this intuition, we assume that governments agree to use the optimal weak binding, τ_A^W, and we analyze the associated off-schedule incentive constraints. Given our focus on weak-binding equilibria, a government with a low type applies its Nash tariff, and so a positive incentive to cheat is present only for a government with a high type. As before, the incentive to cheat is given as $\Omega(\tau_A^W; H) = W_x(\tau^N(H); H) - W_x(\tau_A^W; H)$. The novel aspect of our analysis here concerns the future value of cooperation. In a weak-binding equilibrium, governments' applied tariffs reveal their current types; hence, at the start of any period t, each government can look back at the applied tariffs in period $t-1$ and infer the type that the other government had in that period. Given that types exhibit imperfect persistence, a government can use this inference to better forecast the evolution of the other government's type. In turn, this inference helps the government forecast the expected discounted welfare loss that would be associated with a breakdown in cooperation. Thus, when a government observes its own type and contemplates cheating in period t, its assessment of the future value of cooperation depends on its type in period t and the type that it infers the other government had in period $t-1$.

Our characterization of the future value of cooperation is facilitated by solving a four-by-four recursive system for the values $V_{\gamma\gamma^*}$, where $\gamma \in \{L, H\}$, $\gamma^* \in \{L, H\}$ and $V_{\gamma\gamma^*}$ represents the expected discounted future value of cooperation as evaluated at the start of period t when governments know that their respective types in period $t-1$ were γ and γ^* but have not yet observed their respective period-t types. The recursive system takes the following form:

$$\begin{aligned} V_{\gamma\gamma^*} &= (1-\lambda)^2[A_{\gamma\gamma^*} + \delta V_{\gamma\gamma^*}] + (1-\lambda)\lambda[A_{\gamma\widehat{\gamma}^*} + \delta V_{\gamma\widehat{\gamma}^*}] \\ &\quad + \lambda(1-\lambda)[A_{\widehat{\gamma}\gamma^*} + \delta V_{\widehat{\gamma}\gamma^*}] + \lambda^2[A_{\widehat{\gamma}\widehat{\gamma}^*} + \delta V_{\widehat{\gamma}\widehat{\gamma}^*}] \end{aligned} \tag{6.1}$$

where $\widehat{\gamma} \in \{L, H\}$ and $\widehat{\gamma} \neq \gamma$, $\widehat{\gamma}^* \in \{L, H\}$ and $\widehat{\gamma}^* \neq \gamma^*$, and as discussed next the four "A" variables reflect the per-period value of cooperation under each possible state.

In a weak-binding equilibrium in which governments use the optimal weak binding, along the equilibrium path a government with a low type applies its Nash tariff, $\tau^N(L)$, whereas a government with a high type applies the optimal weak binding, τ_A^W. But if cooperation were to break down, then each government of each type would apply its Nash tariff. Thus, we may define the four "A" variables as follows:

$$\begin{aligned} A_{LL} &= 0, \\ A_{HH} &= J(\tau_A^W; H) - J(\tau^N(H); H) > 0, \\ A_{LH} &= W_x^*(\tau_A^W) - W_x^*(\tau^N(H)) > 0, \\ A_{HL} &= W_x(\tau_A^W; H) - W_x(\tau^N(H); H) < 0 \end{aligned} \tag{6.2}$$

Intuitively, when both governments have the low type, they apply their Nash tariffs whether or not cooperation occurs, and so $A_{LL} = 0$. Similarly, if both governments have the high type, then they both apply τ_A^W under cooperation and $\tau^N(H)$ when cooperation breaks down, and so A_{HH} can be described in terms of joint welfare and is positive. If a government has a low type while its trading partner has a high type, then the government values cooperation, since it applies the same Nash tariff either way but faces a higher tariff on its export good when cooperation fails. It follows that $A_{LH} > 0$. Finally, if a government has a high type and its trading partner does not, then the government actually enjoys a breakdown in cooperation: $A_{HL} < 0$. Intuitively, the government then enjoys applying its Nash tariff when cooperation fails and faces the same tariff on its exports whether or not cooperation fails. Finally, it is apparent from (6.2) that $A_{HH} = A_{LH} + A_{HL}$, and thus that $A_{LH} > A_{HH}$.

Using $A_{LL} = 0$ and $A_{HH} = A_{LH} + A_{HL}$, we may solve (6.1) to get

$$\begin{aligned} V_{LL} &= A_{HH} \frac{\lambda}{(1-\delta)(1-\delta(1-2\lambda))}, \\ V_{HH} &= A_{HH} \frac{1-\lambda-\delta(1-2\lambda)}{(1-\delta)(1-\delta(1-2\lambda))}, \\ V_{LH} &= \frac{A_{HH}}{2(1-\delta)} - \frac{1}{2}\frac{(1-2\lambda)(A_{HL} - A_{LH})}{1-\delta(1-2\lambda)}, \\ V_{HL} &= \frac{A_{HH}}{2(1-\delta)} + \frac{1}{2}\frac{(1-2\lambda)(A_{HL} - A_{LH})}{1-\delta(1-2\lambda)}. \end{aligned} \tag{6.3}$$

Using (6.2), the solutions in (6.3) may be expressed in terms of the underlying welfare functions. Notice that all four values are equal in the limiting case where shocks are transitory:

$V_{\gamma\gamma^*} \equiv A_{HH}/[2(1-\delta)]$ when $\lambda = 1/2$. Under our assumption that $\lambda \in (0, 1/2)$, the following relationships can be established:

$$V_{LH} > V_{HH} > V_{LL} > V_{HL} \tag{6.4}$$

Thus, the expected future value of cooperation is highest when in the previous period a government had a low type while its trading partner had a high type. Intuitively, in that case, if types persist and cooperation occurs, the government already applies its optimal tariff and also faces a below-Nash tariff from its trading partner. At the other extreme, the expected future value of cooperation is lowest when in the previous period a government had the high type and the other government had the low type. Indeed, when λ is sufficiently small so that types are unlikely to change across periods, $V_{HL} < 0$. Further, for all $\lambda \in (0, 1/2)$, $V_{LH} > V_{HH} > V_{LL} > 0$.

We are now ready to consider the off-schedule constraint. As noted, in any period t, a government has a positive short-term incentive to cheat if and only if it has a high type. In period t, the government also has inferred the type that its trading partner had in period $t-1$. Now, the tariff that the trading partner applies in period t is independent of whether or not the government cheats in period t. The government thus understands that a decision to cheat in period t affects the applied tariff of its trading partner starting in period $t+1$. The government that is contemplating cheating in period t thus must consider the future value of cooperation that cheating would sacrifice starting in period $t+1$. Due to persistence, this value is affected by the government's inference as to its trading partner's type in period $t-1$.

For a government with a high type in period t that is contemplating cheating, there are thus two cases. First, if the trading partner applied τ_A^W in period $t-1$ and thereby revealed a high type in that period, then at the close of period t the trading partner would have the low type with probability λ and the high type with probability $1-\lambda$. Hence, in this case, a government in period t regards the future value of cooperation starting in period $t+1$ to be $\delta[\lambda V_{HL} + (1-\lambda)V_{HH}]$. Second, if the trading partner applied $\tau^N(L)$ in period $t-1$ and thus revealed a low type in that period, then at the close of period t the trading partner would have a high type with probability λ and a low type with probability $1-\lambda$. For this case, a government in period t regards the future value of cooperation starting in period $t+1$ to be $\delta[\lambda V_{HH} + (1-\lambda)V_{HL}]$. The off-schedule constraint in period t for a government with a high type is thus $\Omega(\tau_A^W; H) \leq \delta \min\{\lambda V_{HL} + (1-\lambda)V_{HH}, \lambda V_{HH} + (1-\lambda)V_{HL}\}$. Now, given $\lambda \in (0, 1/2)$ and the rankings in (6.4), we may confirm that the minimum is achieved by the second argument: $\lambda V_{HL} + (1-\lambda)V_{HH} > \lambda V_{HH} + (1-\lambda)V_{HL}$. Thus, we may conclude that a weak-binding equilibrium at the optimal weak binding exists for the dynamic game with imperfectly persistent shocks if and only if

$$\Omega(\tau_A^W; H) \leq \delta[\lambda V_{HH} + (1-\lambda)V_{HL}] \equiv RHS(\lambda, \delta). \tag{6.5}$$

Our goal now is to characterize the values of λ and δ for which (6.5) holds. This characterization is complicated by the fact that V_{HL} and V_{HH} are functions of λ and δ, as (6.3) confirms.

We now derive an expression for $RHS(\lambda,\delta)$ and analyze how $RHS(\lambda,\delta)$ is affected by changes in λ. Using (6.3), we find that

$$RHS(\lambda,\delta)=\delta\{\frac{A_{LH}\lambda[2(1-\lambda)-\delta(1-2\lambda)]+A_{HL}[1-\lambda-\delta(1-2\lambda)}{(1-\delta)(1-\delta(1-2\lambda))}\}. \tag{6.6}$$

Using (6.2) and (6.6), we may confirm that

$$RHS(0,\delta)=\frac{\delta}{1-\delta}A_{HL}<0<\frac{\delta}{1-\delta}\frac{A_{HH}}{2}=RHS(\frac{1}{2},\delta). \tag{6.7}$$

Straightforward calculations also reveal that $RHS(\lambda,\delta)$ is strictly increasing in λ over the relevant range: For all $\lambda\in[0,1/2]$,

$$\frac{dRHS(\lambda,\delta)}{d\lambda}=\delta\{\frac{A_{LH}(1-2\lambda)(2-\delta(1-2\lambda))-A_{HL}}{(1-\delta(1-2\lambda))^2}\}>0. \tag{6.8}$$

Thus, as (6.7) confirms, when a government has a high type and infers that its trading partner's type was low in the previous period, then the value of cooperation for this government is negative when λ is sufficiently close to zero. This result parallels our findings above for the dynamic game with perfectly persistent shocks, which corresponds to the limiting case in which $\lambda=0$. Similarly, if λ is sufficiently close to 1/2, then this government would perceive the value of cooperation to be positive. This results parallels our findings above for the repeated game with transitory shocks, which corresponds to the case in which $\lambda=1/2$. Finally, as (6.8) indicates, the value of cooperation is strictly increasing as types become less persistent.[25]

An implication of (6.7) and (6.8) is that, for all $\delta\in(0,1)$ there exists $\lambda^*\in(0,1/2)$ such that $RHS(\lambda^*,\delta)=0$. To characterize this value, we observe that λ^* must satisfy $h(\lambda,\delta)=0$ where $h(\lambda,\delta)\equiv A_{LH}\lambda[2(1-\lambda)-\delta(1-2\lambda)]+A_{HL}[1-\lambda-\delta(1-2\lambda)]$ is the numerator of the fraction in (6.6). We may verify that $h_\lambda(\lambda,\delta)>0$ for all $\lambda\in[0,1/2]$ and $\delta\in[0,1]$.[26] We find as well that $h_\delta(\lambda,\delta)>0$ for all $\lambda\in[0,1/2)$ and $\delta\in[0,1]$, where $h_\delta(\lambda,\delta)=0$ when $\lambda=1/2$.[27] Since $\lambda^*<1/2$, these findings enable us to conclude that

$$\frac{\partial\lambda^*}{\partial\delta}=-\frac{h_\delta(\lambda^*,\delta)}{h_\lambda(\lambda^*,\delta)}<0.$$

Thus, as δ increases, the value of λ at which $RHS(\lambda,\delta)=0$ decreases. Since $h(\lambda,1)=\lambda A_{HH}$, it is clear that, for δ close to unity, λ^* must be close to zero. Likewise, since $h(\lambda,0)=(1-$

[25] We can also show that $RHS(\lambda,\delta)$ is strictly concave in λ, for all $\lambda\in[0,1/2]$ and $\delta\in(0,1)$.

[26] Observe that $h_\lambda(\lambda,\delta)=-4\lambda A_{LH}(1-\delta)+A_{HH}+(1-\delta)(A_{LH}-2A_{HL})\geq h_\lambda(1/2,\delta)=A_{HL}(2\delta-1)+A_{LH}\delta$. Since $A_{HL}<0$, $h_\lambda(1/2,0)>0$. Notice also that $h_\lambda(1/2,1)=A_{HH}>0$. Since $h_\lambda(1/2,\delta)$ is linear in δ, we conclude that $h_\lambda(1/2,\delta)>0$ for all $\delta\in[0,1]$, and it thus follows that $h_\lambda(\lambda,\delta)>0$ for all $\lambda\in[0,1/2]$ and $\delta\in[0,1]$.

[27] We find that $h_\delta(\lambda,\delta)=(2\lambda-1)(A_{LH}\lambda+A_{HL})$; thus, $h_\delta(0,\delta)=-A_{HL}>0$ and $h_\delta(1/2,\delta)=0$. We observe further that $h_{\lambda\lambda\delta}(\lambda,\delta)=4A_{LH}>0$. Thus, if we could show that $h_{\delta\lambda}(1/2,\delta)<0$, then we would have that $h_{\delta\lambda}(\lambda,\delta)<0$ for all $\lambda\in[0,1/2]$. Given that $h_\delta(0,\delta)>0=h_\delta(1/2,\delta)$, we could then conclude that $h_\delta(\lambda,\delta)>0$ for all $\lambda\in[0,1/2)$. We find that $h_{\delta\lambda}(1/2,\delta)=A_{LH}+2A_{HL}$. Substituting and simplifying, we can show that $A_{LH}+2A_{HL}=-(\frac{9(\frac{7}{4}-H)}{(17-2H)(25-4H)})^2<0$.

$\lambda)[2\lambda A_{LH} + A_{HL}]$, as δ gets close to zero, λ^* must be close to the value $\frac{-A_{HL}}{2A_{LH}} < 1/2$, where the inequality follows from $A_{HH} > 0$. In sum, as δ increases from small values near zero to higher values close to unity, λ^* follows a strictly decreasing trajectory from values near $\frac{-A_{HL}}{2A_{LH}} < 1/2$ to values near zero.

We next consider how $RHS(\lambda, \delta)$ is affected by changes in δ. Given we have defined $h(\lambda, \delta)$ so that $RHS(\lambda, \delta) \equiv (\frac{\delta}{1-\delta})(\frac{h(\lambda,\delta)}{(1-\delta(1-2\lambda))})$, we can derive that

$$\frac{\partial RHS(\lambda, \delta)}{\partial \delta} > 0 \text{ for all } \lambda \in [\lambda^*, 1/2]. \tag{6.9}$$

The strict inequality in (6.9) follows from our findings above, where we establish that: (i) $h_\delta(\lambda, \delta) > 0$ for all $\lambda \in [0, 1/2)$ and $h_\delta(1/2, \delta) = 0$; and (ii) $h(\lambda, \delta) > 0$ for $\lambda \in (\lambda^*, 1/2]$ and $h(\lambda^*, \delta) \equiv 0$ where $\lambda^* \in (0, 1/2)$ for all $\delta \in [0, 1)$. We conclude that $RHS(\lambda, \delta)$ is strictly increasing in δ if $\lambda \geq \lambda^*$. For lower values of λ, $RHS(\lambda, \delta)$ may not increase with δ. For example, we know from (6.7) that $RHS(0, \delta) = \frac{\delta}{1-\delta} A_{HL} < 0$, and so $RHS(0, \delta)$ strictly decreases in δ. It is straightforward to show that $\frac{\partial^2 RHS(\lambda,\delta)}{\partial\lambda\partial\delta} > 0$ for all $\lambda \in [0, 1/2]$. For all $\delta \in [0, 1)$, it follows that there exists a critical value $\widetilde{\lambda} \in (0, \lambda^*)$ such that $\frac{\partial RHS(\lambda,\delta)}{\partial\delta} = 0$ for $\lambda = \widetilde{\lambda}$ and $RHS(\lambda, \delta)$ is strictly increasing (strictly decreasing) in δ for $\lambda \in (\widetilde{\lambda}, 1/2]$ (for $\lambda \in [0, \widetilde{\lambda})$).

We now characterize the critical discount factor, $\delta_A^W(\lambda)$, above which governments can support a weak-binding equilibrium in which they apply the tariffs that are induced by the optimal weak binding, τ_A^W. The critical discount factor is the value for δ at which (6.5) binds:

$$\Omega(\tau_A^W; H) = \delta[\lambda V_{HH} + (1-\lambda)V_{HL}] \equiv RHS(\lambda, \delta). \tag{6.10}$$

Let us now fix $\lambda \in (0, 1/2)$ and suppose that we can find a critical discount factor, $\delta_A^W(\lambda)$, at which (6.10) holds. Since the incentive to cheat is strictly positive, it is thus necessary that the critical discount factor is such that $RHS(\lambda, \delta) > 0$. In other words, the critical discount factor must generate a value for λ^* such that $\lambda > \lambda^*$. Thus, if a critical discount factor exists, then we may differentiate (6.10) to find that the critical discount factor must satisfy

$$\frac{\partial \delta_A^W(\lambda)}{\partial \lambda} = -\frac{\partial RHS(\lambda, \delta)/\partial\lambda}{\partial RHS(\lambda, \delta)/\partial\delta} < 0, \tag{6.11}$$

where the inequality follows from (6.8), (6.9) and $\lambda > \lambda^*$. We thus conclude that, if a critical discount factor exists, then it is strictly lower when governments' types are less persistent. Put differently, it is easier to support a weak-binding equilibrium when types are more transitory.

The remaining step is to address the existence of the critical discount factor. Our arguments are most easily developed with reference to Figure 9.[28] The horizontal line represents the

[28] Figure 9 embodies the properties of $RHS(\lambda, \delta)$ derived above. In particular, $RHS(\lambda, \delta)$ is strictly increasing in λ, strictly increasing in δ for $\lambda \geq \lambda^*(\delta)$, and more positively sloped with respect to λ when δ is raised. Notice as well that the function $\lambda^*(\delta)$ is depicted as falling when δ rises. To make the graph less cluttered, we have not explicitly labeled $\widetilde{\lambda}$.

incentive to cheat for a government with a high type. To begin, we consider the limiting case where $\lambda = 1/2 \equiv \lambda_0$. Given the properties of $RHS(\lambda, \delta)$ derived above, we know that there exists some value $\delta_0 \in (0, 1)$ such that $RHS(\lambda_0, \delta_0) = \Omega(\tau_A^W; H)$ Thus, δ_0 is the critical discount factor for the limiting case in which $\lambda = 1/2$. Since in this case shocks are transitory, we know that δ_0 is in fact the critical discount factor derived in (4.7) when $\eta_H = 1/2$ for the repeated game with transitory shocks. Now suppose that we lower λ to the value $\lambda_1 \in (0, 1/2)$. As shown in Figure 9, using the properties derived above we can find a new critical discount factor, δ_1, such that $RHS(\lambda_1, \delta_1) = \Omega(\tau_A^W; H)$. Consistent with (6.11), we see that the critical discount factor is higher as λ is lowered: $\delta_1 \in (\delta_0, 1)$. Continuing this process one more time, we may lower λ from λ_1 to λ_2 and thereby determine a new critical discount factor δ_2, where $\delta_2 \in (\delta_1, 1)$. Finally, as λ gets close to zero, the critical discount factor must get close to unity (so that the RHS curve in Figure 9 becomes almost vertical). In this general way, we may construct the critical discount factor, $\delta_A^W(\lambda)$, as a function of λ.

We may now summarize our findings in the following proposition:

Proposition 11: *Consider the dynamic game with imperfectly persistent shocks. For all* $\lambda \in (0, 1/2)$, t*here exists* $\delta_A^W(\lambda) \in (0, 1)$ *such that (i) for all* $\delta \geq \delta_A^W(\lambda)$*, there exists a weak-binding equilibrium in which the weak binding is set at the optimal weak binding,* τ_A^W*, and (ii) for all* $\delta < \delta_A^W(\lambda)$*, there does not exist a weak-binding equilibrium in which the weak binding is set at the optimal weak binding,* τ_A^W*. The critical discount factor,* $\delta_A^W(\lambda)$*, is strictly decreasing in* λ*, approaches unity as* λ *approaches zero, and approaches the critical discount factor defined in (4.7) when* $\eta_H = 1/2$ *for the repeated game with transitory shocks as* λ *approaches* $1/2$.

As Proposition 11 indicates, governments can enforce a weak-binding equilibrium in which the weak binding is set at the optimal weak binding if they are sufficiently patient, where greater patience is required when types are more persistent. In particular, a government with a low type can then apply its Nash tariff and reveal that it is weak without inducing the other government to cheat in the following period, if the latter government takes the long view and recognizes that the former government will one day be strong again. When λ is close to $1/2$, types are essentially transitory, and our findings are similar to those derived earlier for the repeated game with transitory shocks. Likewise, when λ is close to zero, types are essentially permanent, and our findings are analogous to those derived above for the dynamic game with perfectly persistent shocks. Indeed, when λ is close to zero, the critical discount factor is close to unity; thus, in this case, governments can enforce a weak-binding equilibrium in which the weak binding is set at the optimal weak binding if and only if they have approximately unlimited patience.

As Proposition 11 confirms, for a given $\lambda \in (0, 1/2)$, if the discount factor falls below $\delta_A^W(\lambda)$, then there does not exist a weak-binding equilibrium that uses the optimal weak binding. As in Proposition 7 for the repeated game with transitory shocks, we may then construct a weak-binding equilibrium in which the weak binding is set at a level that is above the optimal weak binding. Intuitively, and as before, a higher bound tariff rate reduces the incentive to cheat for a government with a high type, since the higher bound enables such a government to apply a tariff closer to its Nash tariff. Similarly, if we fix any level of patience $\delta \in (0, 1)$ and increase the level of persistence (i.e., lower the level of λ), then we encounter a critical level of persistence such that governments can enforce a weak-binding equilibrium only if they set the weak binding strictly above the optimal weak binding. The divergence between the bound tariff and the applied tariff of a government with a low type increases as the tariff binding is raised. The analysis here thus leads to the following prediction: when types are more persistent, larger and more persistent divergences between applied and bound tariffs may be observed.

For a fixed level of patience, if types are very persistent, even the "most-cooperative" weak-binding equilibrium offers only a modest improvement on the Nash equilibrium. Intuitively, in this case, when a government has a high type and infers that its trading partner's type was low in the previous period, this government cannot be deterred from cheating unless the bound tariff is sufficiently close to its Nash tariff. This observation captures one sense in which the enforcement problem highlighted in our analysis of the dynamic game with perfectly persistent shocks exerts an influence as well in the dynamic game with imperfectly persistent shocks. If governments were to face such a situation, it is conceivable that they might attempt to enforce a pooling equilibrium in which the applied tariff always equals the bound rate. Their behavior would then be analagous to that described in Proposition 10 for the dynamic game with perfectly persistent types.

7. Gradualism and Dynamic Screening

In this section, we consider again the dynamic game with perfectly persistent shocks; however, we now consider a new class of equilibria, in which governments reach successive agreements over time with regard to bound tariff rates. In particular, we allow that the tariffs applied by governments in the course of satisfying the weak binding of the first agreement may affect the weak bindings that governments negotiate in the second agreement. We thus refer to such equilibria as "history-dependent weak-binding equilibria." Allowing for any discount factor, we construct an equilibrium of this kind that offers higher-than-Nash welfare for both governments for every possible pair of types that governments may have. The equilibrium entails dynamic screening and is suggestive of a new information-based theory of gradualism in trade agreements.

We also consider the dynamic game with perfectly persistent shocks in Section 5. We observe there that weak-binding equilibria partition the public history into "good" and "bad" histories and thus allow for limited history dependence. We now allow for a small degree of additional history dependence. To this end, we consider the possibility of an equilibrium that can be described in terms of a weak binding for period one, τ_1^W, and a function, $\tau_2^W(\tau_1, \tau_1^*)$, that determines the weak binding for a government in periods $t = 2, ..., \infty$ as a function of the tariffs τ_1 and τ_1^* that this government and its trading partner respectively apply in period one, whenever $\tau_1 \leq \tau_1^W$ and $\tau_1^* \leq \tau_1^W$. We refer to τ_1^W as the *first-period weak binding*. For $\tau_1 \leq \tau_1^W$ and $\tau_1^* \leq \tau_1^W$, we refer to $\tau_2^W(\tau_1, \tau_1^*)$ as a government's *induced second-period weak binding*, where it is understood that the "second period" weak binding is used in all periods $t \geq 2$.

As before, if a government applies a tariff in excess of its weak binding, then in all future periods governments abandon cooperation and apply Nash tariffs. Thus, if either government's first-period applied tariff exceeds the first-period weak binding, τ_1^W, then governments revert to Nash play in the second and all subsequent periods; likewise, a reversion to Nash play commences in period $t+1$ if in any period $t \geq 2$ a government applies a tariff in excess of its induced second-period weak binding, $\tau_2^W(\tau_1, \tau_1^*)$.

We may interpret the first-period and induced second-period weak bindings as corresponding to two trade agreements. The first agreement is negotiated at date zero and determines a common weak binding, τ_1^W, for both governments in period one. The second agreement is negotiated at the end of period one and thus after the public observation of period-one applied tariffs. If governments apply different tariffs in period one, then they may negotiate different weak bindings in the period-two agreement. In particular, if governments have different types, then it is possible that they both satisfy the weak-binding requirement in period one and yet apply different tariffs in that period. While the definition of a weak-binding equilibrium would require that future applied tariffs are independent of the particular ways in which governments satisfy their first-period weak binding, we introduce here some additional history dependence and allow that the induced second-period weak bindings may reflect the particular ways in which governments satisfy their first-period weak binding. In other words, different "good" histories may now be associated with different induced weak bindings in the second agreement.

For the dynamic game with perfectly persistent shocks, we define a *history-dependent weak-binding equilibrium* as an equilibrium in which there exists τ_1^W and $\tau_2^W(\tau_1, \tau_1^*)$ such that (i) along the equilibrium path, in period one governments apply tariffs that are equal to or below their first-period weak binding, τ_1^W, and in any period $t \geq 2$ governments apply tariffs that are equal to or below their respective induced second-period weak bindings, $\tau_2^W(\tau_1, \tau_1^*)$ and $\tau_2^W(\tau_1^*, \tau_1)$, and (ii) for any $\tau_1 \leq \tau_1^W$ and $\tau_1^* \leq \tau_1^W$ and $t \geq 3$, the tariffs that governments apply in period t are independent of the public history of applied tariffs in periods $\widetilde{t} = 2, ..., t-1$, so long as no government has applied a tariff in any period $\widetilde{t}$ in excess of its induced second-period weak binding.

With the notion of a history-dependent weak-binding equilibrium, we thus allow that first-period tariffs affect the determination of the induced second-period weak bindings and thereby the tariffs that are applied in future periods. In effect, starting the second period, government behavior must be described by a weak-binding equilibrium, once that concept is generalized slightly to allow for asymmetric weak bindings. But the second-period weak binding itself represents a new channel for history dependence.

To understand the potential value of a history-dependent weak-binding equilibrium, it is helpful to recall the limited scope for cooperation in a weak-binding equilibrium when types are perfectly persistent. In such an equilibrium, a single weak binding, τ^W, is used, regardless of the governments' actual types. For $\tau^W \in (\tau^N(L), \tau^N(H))$, a government with a low type is not constrained by the weak binding; indeed, it would apply its optimal tariff, $\tau^N(L)$, whether or not its partner had previously violated the agreement. A government with a low type is thus unable to retaliate in an effective way. This leads to an enforcement problem in period two, if the government's period-one applied tariff reveals its type and the trading partner has a high type. As established in Proposition 8, a weak-binding equilibrium with $\tau^W \in (\tau^N(L), \tau^N(H))$ thus fails to exist. Suppose now, though, that governments have the opportunity to negotiate a new agreement in period two. They might then set a weak binding below $\tau^N(L)$ for a government that has revealed itself to have a low type and a different weak binding below $\tau^N(H)$ for a government that has revealed itself to be a high type. Starting in period two, each government would then have the capacity to retaliate in an effective manner, since a government's bound tariff is below its optimal tariff, regardless of the government's type.

This intuition could be formalized in several ways. Here, we construct a simple history-dependent weak-binding equilibrium in which governments achieve greater-than-Nash welfare regardless of their types and for any positive discount factor. In the constructed equilibrium, the first-period weak binding is $\tau_1^W = \tau^N(H)$. The induced second-period weak binding for a government is then $\tau^N(H) - \Delta_H$ if this government's first-period applied tariff is $\tau^N(H)$, where $\Delta_H > 0$. If instead a government's first-period applied tariff is strictly below $\tau^N(H)$, then the induced second-period weak binding for this government is $\tau^N(L) - \Delta_L$, where $\Delta_L > 0$. The proposed system of weak bindings is simple, in that a government's induced second-period weak binding depends on its applied first-period tariff and not on that of its trading partner.

The next step is to determine the applied tariffs that governments select, when such a binding system is present. Consider a government with the high type. If Δ_H is sufficiently small, then such a government's best option in the first period is to apply its Nash tariff, $\tau^N(H)$. By doing so, the government achieves its optimal tariff in period one and induces a second-period weak binding that is only slightly below its Nash tariff.[29] For Δ_H sufficiently small, the government's best choice is to apply a tariff equal to the induced second-period weak binding, $\tau^N(H) - \Delta_H$, in all future periods. Intuitively, the gain from cheating to $\tau^N(H)$ is "second-order," since $W_x(\tau; H)$ is maximized at $\tau = \tau^N(H)$; moreover, regardless of the trading partner's type, the future welfare cost to a government of cheating and thereby triggering a reversion to Nash tariffs is "first-order," since the government's welfare declines strictly when

[29]A government with a high type could violate its binding and apply a first-period tariff in excess of $\tau^N(H)$; however, it would then enjoy less welfare in the first period and also in later periods, since its trading partner would apply Nash tariffs in all subsequent periods. Similarly, it could apply a first-period tariff below $\tau^N(H)$, but it would then enjoy less welfare in the first period and induce a lower second-period weak binding, $\tau^N(L) - \Delta_L$.

the trading partner's tariff increases by $\Delta_H > 0$ or $\Delta_L > 0$. Notice in particular that even a trading partner with a low type can now retaliate effectively, because its induced second-period weak binding, $\tau^N(L) - \Delta_L$, is below its Nash tariff. When Δ_L is sufficiently small, a similar argument is available for a government with a low type.

We may now state the following proposition:

Proposition 12: *Consider the dynamic game with perfectly persistent shocks. For $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small, there exists history-dependent weak-binding equilibrium with a first-period weak binding $\tau_1^W = \tau^N(H)$ and an induced second-period weak binding $\tau_2^W(\tau_1, \tau_1^*)$ such that, for $\tau_1^* \leq \tau^N(H)$, $\tau_2^W(\tau_1, \tau_1^*) = \tau^N(H) - \Delta_H$ when $\tau_1 = \tau^N(H)$ and $\tau_2^W(\tau_1, \tau_1^*) = \tau^N(L) - \Delta_L$ when $\tau_1 < \tau^N(H)$. Along the equilibrium path, for $\gamma \in \{L, H\}$, a government with type γ applies the tariff $\tau^N(\gamma)$ in period one and the tariff $\tau^N(\gamma) - \Delta_\gamma$ in period two and all subsequent periods. For all possible types that governments may realize, each government earns strictly higher discounted welfare in the constructed history-dependent weak-binding equilibrium than they would were they to apply Nash tariffs in all periods.*

The remaining details of the proof are provided in the Appendix.

Proposition 12 captures a form of dynamic screening. With their first-period applied tariffs, governments reveal their types in period one. They then agree upon new weak bindings in period two which reflect this information. The constructed equilibrium can be interpreted as a form of gradualism in trade agreements.[30] The first-period agreement uses a single and high weak binding. After observing the tariffs that are applied under this agreement, and making the appropriate inferences, governments then negotiate a new agreement in period two in which the weak bindings are further lowered. The negotiated reduction in the weak binding is greater if a government has a low type and thus applies a first period tariff that is strictly below the first-period weak binding. All governments apply their bound tariffs in the second period.

Finally, we emphasize that Proposition 12 does not chararacterize the most-cooperative history-dependent weak-binding equilibrium. We leave this important task to future work. Instead, our main goal here is to show that the introduction of a modest degree of history dependence can counter the ratchet effect and restore cooperation in the dynamic game with perfectly persistent shocks. The constructed equilibrium is also suggestive of a new information-based theory of gradualism in trade agreements.

8. Conclusion

We consider the design of self-enforcing trade agreements among governments with private information about the political pressures that they face. By limiting tariffs, a trade agreement

[30] An alternative interpretation is that governments have no agreement in period one and apply their tariffs with an understanding that their period-one choices may impact the bindings that are negotiated in period two as part of a trade agreement.

can reduce the standard inefficiencies that are associated with the terms-of-trade externality. At the same time, if a trade agreement allows governments some discretion when setting tariffs, it can facilitate better matching between a government's applied tariffs and its political pressures. To explore these roles for trade agreements, we examine a static model and a sequence of dynamic models. The dynamic models are distinguished by the extent to which the private information that governments possess is persistent through time. In this context, we argue that an agreement to use a weak binding offers important advantages over other alternatives. We note, though, that governments may hesitate to fully exploit the benefitical "downward discretion" that such an agreement allows, if private information is persistent relative to the patience of governments. In particular, if a government is perceived as "weak," then it cannot threaten effective retaliation; thus, an agreement to apply the tariffs induced by the optimal weak binding may not be self-enforcing. We also consider some ways in which governments might respond to the associated "ratchet effect." Our analysis suggests a new information-based theory of gradualism in trade agreements.

Several important extensions await. First, we consider a model in which each country imports one good and exports another good, whereas in reality countries import and export many goods. Most of our results extend to the multi-good context in a straightforward fashion.[31] In some cases, though, the multi-good extension introduces new and interesting considerations. In particular, it would be interesting to examine the possibility of pooling equilibria in the dynamic game with perfectly persistent shocks, when multiple goods are traded. When governments attempt to use the applied tariffs induced by the optimal weak binding, a ratchet effect again arises if a government has a low type for a sufficient number of its goods; however, in the multi-good case, it may be possible that a government could hide its weakness to a sufficient degree, by pooling and applying the bound tariff rate for many but not all products. Second, in the two-type framework analyzed in this paper, there is little scope for analyzing trade agreements with weak bindings (downward discretion) *and* safeguard rules (upward discretion). As Bagwell and Staiger (2005) show, in a model with a continuum of transitory types, it is possible to analyze weak bindings and safeguard rules simultaneously. Important future work might consider a model with more than two types where the types exhibit some persistence over time. Such a framework might also offer insight with respect to WTO remedies (countermeasures), which occur along the equilibrium path when one government exercises upward discretion. Finally, an important task for future work is to characterize and interpret the optimal form of gradualism in self-enforcing trade agreements among governments with persistent private information.

[31] We can show, for example, that the critical discount factor for the repeated game with transitory shocks is unaltered when the model is replicated to allow for K identical import goods and K identical export goods, where $K \geq 1$. A government is most tempted to cheat when it has the high type on all K goods at once. In any period, a government expects that its trading partner will have the high (low) type on $\eta_H K$ ($\eta_L K$) goods. It is now easy to see that K cancels from the relevant off-schedule constraint.

9. References

Atkeson, A., S. Athey, P. Kehoe (2005), "The Optimal Degree of Monetary Policy Discretion," *Econometrica*, 73.5, 1431-76.

Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information," *Rand Journal of Economics*,32, 428-65.

Athey, S. and K. Bagwell (2006), "Collusion with Persistent Cost Shocks," mimeo, Columbia University.

Athey, S., K. Bagwell and C. Sanchirico (2004), "Collusion and Price Rigidity," *Review of Economic Studies*, 71: 317-49.

Bagwell, K. (forthcoming), "Remedies in the WTO: An Economic Perspective," in Merit E. Janow, Victoria Donaldson, and Alan Yanovich (eds.), **The WTO: Governance, Dispute Settlement and Developing Countries**, Juris Publishing, Inc..

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

Bagwell, K. and R. W. Staiger (1999), "An Economic Theory of GATT," *American Economic Review*, 89, 215-48.

Bagwell, K. and R. W. Staiger (2001), "Reciprocity, Non-discrimination and Preferential Agreements in the Multilateral Trading System," *European Journal of Political Economy*, 17, 281-325.

Bagwell, K. and R. W. Staiger (2002), **The Economics of the World Trading System**, Cambridge, MA.: The MIT Press.

Bagwell, K. and R. W. Staiger (2005), "Enforcement, Private Political Pressure and the GATT/WTO Escape Clause," *The Journal of Legal Studies*, 34(2), 471-513.

Bagwell, K. and A. O. Sykes (2004), "Chile - Price Band System and Safeguard Measures Relating to Certain Agriculture Products," *World Trade Review*, 3.3, 507-28.

Baldwin, R. (1987), "Politically Realistic Objective Functions and Trade Policy," *Economic Letters*, 24, 287-90.

Bond, E. and J.-H. Park (2002), "Gradualism in Trade Agreements with Asymmetric Countries," *Review of Economic Studies*, 69, 379-406.

Broda, C., N. Limao, and D. Weinstein (forthcoming), "Optimal Tariffs: The Evidence," *American Economic Review*.

Chisik, R. (2003), "Gradualism in Free Trade Agreements: A Theoretical Justification," *Journal of International Economics*,59.2, 367-97.

Devereux, M. (1997), "Growth, Specialization and Trade Liberalization," *International Economic Review*, 38, 565-85.

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

Feenstra, R. and T. Lewis (1991), "Negotiated Trade Restrictions with Private Political Pressure," *Quarterly Journal of Economics*, 196, 1287-1307.

Freixas, X., R. Guesnerie and J. Tirole (1985), "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52, 173-92.

Fudenberg, D., D, Levine and E. Maskin (1994), "The Folk Theorem with Imperfect Public Monitoring," *Econometrica*, 62, 997-1039.

Fudenberg and Tirole (1991), **Game Theory**, Cambridge, MA.: The MIT Press.

Grossman, G. M. and E. Helpman (1994), "Protection for Sale," *American Economic Review*, 84, 833-50.

Grossman, G. M. and E. Helpman (1995), "Trade Wars and Trade Talks," *Journal of Political Economy*, 103, 675-708.

Horn, H., G. Maggi and R. W. Staiger (2006), "Trade Agreements As Endogenously Incomplete Contracts," mimeo, Stanford University.

Laffont, J.-J. and J. Tirole (1988), "The Dynamics of Incentive Contracts," *Econometrica*, 56, 1153-75.

Lee, G. M. (2007), "Trade Agreements with Domestic Policies As Disguised Protection," *Journal of International Economics*, 71.1, 241-59.

Maggi, G. (1999), "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, 89, 190-214.

Maggi, G. and A. Rodriguez-Clare (2007), "A Political-Economy Theory of Trade Agreements," *American Economic Review*, 97.4, 1374-1406.

Martin, A. and W. Vergote (2004), "A Case for Strategic Antidumping," mimeo, Columbia University.

Park, J.-H. (2006), "Private Trigger Strategies in the Presence of Concealed Trade Barriers," mimeo, Seoul National University.

Skryzpacz, A. and H. Hopenhayn (2004), "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory*, 114.1, 153-69.

Staiger, R. W. (1995), "A Theory of Gradualin Trade Liberalization," in A. Deardorff, J. Levinsohn, and R. Stern (eds.), **New Directions in TradeTheory**, Ann Arbor,MI:: University of Michigan Press.

10. Appendix

Proof of Proposition 1: To prove (i), let $\Delta_\tau(\eta_H) \equiv \tau_A^S - E\tau^N(\gamma) = \tau^E(E\gamma) - E\tau^N(\gamma)$. Using (2.3) and (2.5), we find $\Delta_\tau(0) = \tau^E(L) - \tau^N(L) < 0$ and $\Delta_\tau(1) = \tau^E(H) - \tau^N(H) < 0$. Calculations confirm $\Delta_\tau^{''}(\eta_H) > 0$. Thus, for all $\eta_H \in (0,1)$, $\Delta_\tau(\eta_H) < 0$. To prove (ii), let $\Delta_{W_x^*}(\eta_H) \equiv W_x^*(\tau_A^S) - EW_x^*(\tau^N(\gamma)) = W_x^*(\tau^E(E\gamma)) - EW_x^*(\tau^N(\gamma))$. Observe that $\Delta_{W_x^*}(0) = W_x^*(\tau^E(L)) - W_x^*(\tau^N(L)) > 0$ and $\Delta_{W_x^*}(1) = W_x^*(\tau^E(H)) - W_x^*(\tau^N(H)) > 0$. Calculations confirm $\Delta_{W_x^*}^{''}(\eta_H) < 0$. Thus, for all $\eta_H \in (0,1)$, $\Delta_{W_x^*}(\eta_H) > 0$. **QED**

Proof of Proposition 2: We begin by reporting some properties of Δ. First, calculations confirm that

$$\frac{\partial \Delta}{\partial \eta_H} = \frac{4(H-L)}{49}(\tau_A^S - \tau^N(L))(2 + \tau_A^S + \tau^N(L)) + J(\tau^N(L), H) - J(\tau^N(H), H). \quad (10.1)$$

Second, using (10.1), we can confirm that Δ is strictly convex in η_H. Third, if η_H is set at a value such that $\tau_A^S = \tau^N(L)$, then $\Delta = \eta_H \frac{\partial \Delta}{\partial \eta_H}$.

We now prove part (i). Referring to Figure 3, given that H and L are such that $L(H,1) > L > L^N(H)$, they reside in Region A. Since $L(H, \eta_H)$ strictly decreases as η_H rises, there exists $\widetilde{\eta}_H \in (1/3, 1)$ such that $L = L(H, \widetilde{\eta}_H)$. For all $\eta_H \in (\widetilde{\eta}_H, 1)$, $L < L(H, \eta_H)$, and so $\tau_A^S > \tau^N(L)$ is satisfied. If $\eta_H = \widetilde{\eta}_H$, then $\tau_A^S = \tau^N(L)$. Using (10.1), we would then have $\frac{\partial \Delta}{\partial \eta_H} = J(\tau^N(L), H) - J(\tau^N(H), H) > 0$, where the inequality follows since $L > L^N(H)$. Using the third property of Δ, it follows that $\Delta > 0$ when $\eta_H = \widetilde{\eta}_H$; thus, $\Delta > 0$ for all $\eta_H \in (\widetilde{\eta}_H, 1)$.

Consider next part (ii). Referring again to Figure 3, given that H and L are such that $L^N(H) > L$, they reside in Region B. As before, since $L(H, \eta_H)$ strictly decreases as η_H rises, there exists $\widetilde{\eta}_H \in (1/3, 1)$ such that $L = L(H, \widetilde{\eta}_H)$. For all $\eta_H \in (\widetilde{\eta}_H, 1)$, $L < L(H, \eta_H)$, and so $\tau_A^S > \tau^N(L)$ is satisfied. If $\eta_H = \widetilde{\eta}_H$, then $\tau_A^S = \tau^N(L)$. Using (10.1), we would then have $\frac{\partial \Delta}{\partial \eta_H} = J(\tau^N(L), H) - J(\tau^N(H), H) < 0$, where the inequality follows since $L^N(H) > L$. Using the third property of Δ, it follows that $\Delta < 0$ when $\eta_H = \widetilde{\eta}_H$. By the definition of Δ, when $\eta_H = 1$, we have that $\Delta = J(\tau^E(H); H) - J(\tau^N(H); H) > 0$. Since Δ is strictly convex in η_H, there exists a unique value $\widehat{\eta}_H \in (\widetilde{\eta}_H, 1)$ such that for all $\eta_H \in (\widetilde{\eta}_H, \widehat{\eta}_H)$, $\Delta < 0$ while for all $\eta_H \in (\widehat{\eta}_H, 1)$, $\Delta > 0$. **QED**

Proof of Proposition 4: When governments agree to use the optimal weak binding, they apply the tariffs $\tau^N(L)$ and $\tau_A^W = \tau^E(H) < \tau^N(H)$ in the low- and high-type states, respectively. It follows that the expected Nash tariff is strictly higher than the expected applied tariff when governments agree to use the optimal weak binding. To establish the second part of the proposition, we define $\Delta_{\tau^W}(\eta_H) \equiv \eta_H \tau^E(H) + (1-\eta_H)\tau^N(L) - \tau_A^S = \eta_H \tau^E(H) + (1-\eta_H)\tau^N(L) - \tau^E(E\gamma)$. Using (2.3) and (2.5), we find $\Delta_{\tau^W}(0) = \tau^N(L) - \tau^E(L) > 0$ and $\Delta_{\tau^W}(1) = \tau^E(H) - \tau^E(H) = 0$. Calculations confirm $\Delta_\tau''(\eta_H) < 0$. Thus, for all $\eta_H \in (0, 1)$, $\Delta_{\tau^W}(\eta_H) > 0$. **QED**

Proof of Proposition 6: Assume (3.6) holds. We establish in our third claim that the solution to the relaxed program satsifies the inequalities stated in the proposition. We now show that the solution to the relaxed program satisfies the omitted incentive constraint and thus is also the solution to the second-best program. Recall from our first claim that the solution to the relaxed program satsifies $W_x(\tau(L); L) = W_x(\tau(H); L)$ and $\tau(H) > \tau(L)$. Thus,

$$
\begin{aligned}
W_x(\tau(H); H) &= W_x(\tau(H); L) + (H - L) f(\tau(H)) \\
&= W_x(\tau(L); L) + (H - L) f(\tau(H)) \\
&= W_x(\tau(L); H) + (H - L)[f(\tau(H)) - f(\tau(L))] \\
&> W_x(\tau(L); H),
\end{aligned}
$$

where the inequality follows since $H > L$, $\tau(H) > \tau(L)$ and $f' > 0$. The omitted incentive constraint is thus satisfied. Finally, as we establish above and as Figure 4 illustrates, $L(H, 1)$ lies strictly above $L^{FB}(H)$ which in turn lies strictly above $L(H, 1/2)$. It follows that there exists some critical $\widetilde{\eta}_H \in (1/2, 1)$ such that for all $\eta_H \in (\widetilde{\eta}_H, 1)$ there exists values for L and H that satisfy (3.6) and our parameter restriction. **QED**

Proof of Proposition 12: To begin, consider any period $t \geq 2$ and suppose that a government has a high type. First, suppose that the government and its trading partner have both behaved as high types in all previous periods (i.e., each government applied $\tau^N(H)$ in period one and $\tau^N(H) - \Delta_H$ in all subsequent periods). In period t, the most attractive deviation for the government is to cheat by applying the tariff $\tau^N(H)$. It then enjoys a one-period gain of $\Omega(\tau^N(H) - \Delta_H; H)$ and triggers a reversion to Nash tariffs in all future periods. Given that the trading partner is believed to have a high type, the off-schedule constraint can be written as $\Omega(\tau^N(H) - \Delta_H; H) \leq \frac{\delta}{1-\delta}[W_x^*(\tau^N(H) - \Delta_H) - W_x^*(\tau^N(H)) - \Omega(\tau^N(H) - \Delta_H; H)]$, or equivalently

$$
W_x(\tau^N(H); H) - W_x(\tau^N(H) - \Delta_H; H) \leq \delta[W_x^*(\tau^N(H) - \Delta_H) - W_x^*(\tau^N(H))]. \tag{10.2}
$$

The LHS and RHS of (10.2) are zero when $\Delta_H = 0$; further, the derivative of the LHS with respect to Δ_H is zero at $\Delta_H = 0$, while for $\delta > 0$ the derivative of the RHS with respect to Δ_H is positive at $\Delta_H = 0$. We thus conclude that (10.2) holds for $\Delta_H > 0$ sufficiently small.

Second, suppose that the government has behaved as a high type in all previous periods, and that its trading partner has behaved as a low type in all previous periods (i.e., the government applied $\tau^N(H)$ in period one and $\tau^N(H) - \Delta_H$ in all subsequent periods, and the trading partner applied $\tau^N(L)$ in period one and $\tau^N(L) - \Delta_L$ in all subsequent periods). In period t, the most attractive deviation for the government is to cheat by applying the tariff $\tau^N(H)$. It then enjoys a one-period gain of $\Omega(\tau^N(H) - \Delta_H; H)$ and triggers a reversion to Nash tariffs in all future periods. Given that the trading partner is believed to have a low type, the off-schedule constraint can be written as

$$W_x(\tau^N(H); H) - W_x(\tau^N(H) - \Delta_H; H) \leq \delta[W_x^*(\tau^N(L) - \Delta_L) - W_x^*(\tau^N(L))]. \tag{10.3}$$

The LHS and RHS of (10.3) are zero when $\Delta_H = \Delta_L = 0$; further, the derivative of the LHS with respect to Δ_H is zero at $\Delta_H = 0$, while for $\delta > 0$ the derivative of the RHS with respect to Δ_L is positive at $\Delta_L = 0$. We thus conclude that (10.3) holds for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small.

We note that exactly analogous arguments apply for any period $t \geq 2$ when the government has a low type and has behaved accordingly in all previous periods, while its trading partner has a low or high type and has behaved accordingly in all previous periods as well.

For $t \geq 2$, suppose next that a government has a high type and that its partner behaved as a high type in period one (i.e., applied $\tau^N(H)$) but in some later period deviated by applying a tariff below its induced second-period weak binding (i.e., below $\tau^N(H) - \Delta_H$). At this point, play have moved off the equilibrium path, and for simplicity we may assume that the government has "passive beliefs" and continues to believe that its trading partner has a high type. Given this state of play, we may specify the equilibrium strategy so that a trading partner with a high type applies a tariff equal to its induced second-period weak binding in period t and thereafter, provided that the government continues to meet its induced second-period weak binding. Given its belief and understanding of the equilibrium strategy, the relevant off-schedule constraint for the government is again described by (10.2), which we know holds for $\Delta_H > 0$ sufficiently small. A similar argument applies if a government has a high type and its trading partner behaved as a low type in period one (i.e., applied $\tau^N(L)$) but in some later period deviated by applying a tariff below its induced second-period weak binding (i.e., below $\tau^N(L) - \Delta_L$). In this case, we refer to (10.3) and conclude that the government will not deviate for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small. Likewise, for $t \geq 2$, if a government has a low type and its partner behaved as a low (high) type in period one and then later deviated by applying a tariff below its induced second-period weak binding, then the government will not deviate from applying a tariff equal to its induced second-period weak binding, for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small.

For $t \geq 2$, another possibility is that the trading partner applied an equilibrium tariff in period one (either $\tau^N(L)$ or $\tau^N(H)$) and then later deviated by applying a tariff above its induced second-period weak binding. In this case, the government's best choice is to revert to its Nash tariffs, since equilibrium strategies now require that its trading partner applies Nash tariffs in all remaining periods independent of the behavior of the government.

For $t \geq 2$, a final possibility is that the trading partner deviated in period one. If the trading partner applied a first-period tariff in excess of $\tau^N(H)$, then the government anticipates that its trading partner will apply Nash tariffs in all future periods, independent of the behavior of the government. The government's best choice is then to apply its Nash tariff in all future periods. If the trading partner applied a first-period tariff below $\tau^N(H)$ but different from $\tau^N(L)$, then we are off the equilibrium path and may assume that the government believes its trading partner has the low type. Given this state of play, we may specify the equilibrium strategy so that a trading partner with a low type applies a tariff equal to its induced second-period weak binding, $\tau^N(L) - \Delta_L$, in period 2 and thereafter, provided that the government continues to meet its induced second-period weak binding. Given its belief and understanding of the equilibrium strategy, the relevant off-schedule constraint for the government with a high type is again described by (10.3), which we know holds for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small. A similar argument applies if the government has a low type.

We now roll back and consider the first period of play. Regardless of its type, a government applies its Nash tariff in equilibrium in period one; thus, any deviation serves to lower the government's first-period welfare. If the deviation exceeds the first-period weak binding, then it triggers Nash reversion and also serves to lower the government's welfare in future periods, for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small. This follows since the gain to the government of being able to apply its Nash tariffs in future periods is second-order in size while the cost to the government of the application of Nash tariffs by its trading partner is first-order in size. Finally, if the deviation is below the first-period weak binding, then a government with a low type does not alter future play, while a government with a high type induces a lower second-period weak binding then it would have otherwise and thus gains less under cooperation in the future. Such a government might then apply its Nash tariff and cheat in the future; however, in that event, it would have done better applying its Nash tariff in all periods (including $t = 1$), and for $\Delta_H > 0$ and $\Delta_L > 0$ sufficiently small the equilibrium payoffs to such a government are higher yet.

Finally, we note that the constructed equilibrium generates greater discounted welfare for governments, regardless of their types, than they receive under Nash play. Consistent with the discussion above, this follows since, for any pair of types and starting at Nash tariffs, each government enjoys a first-order gain from a small reduction in the tariff of its trading partner and experiences only a second-order loss from a small reduction in its own tariff. **QED**

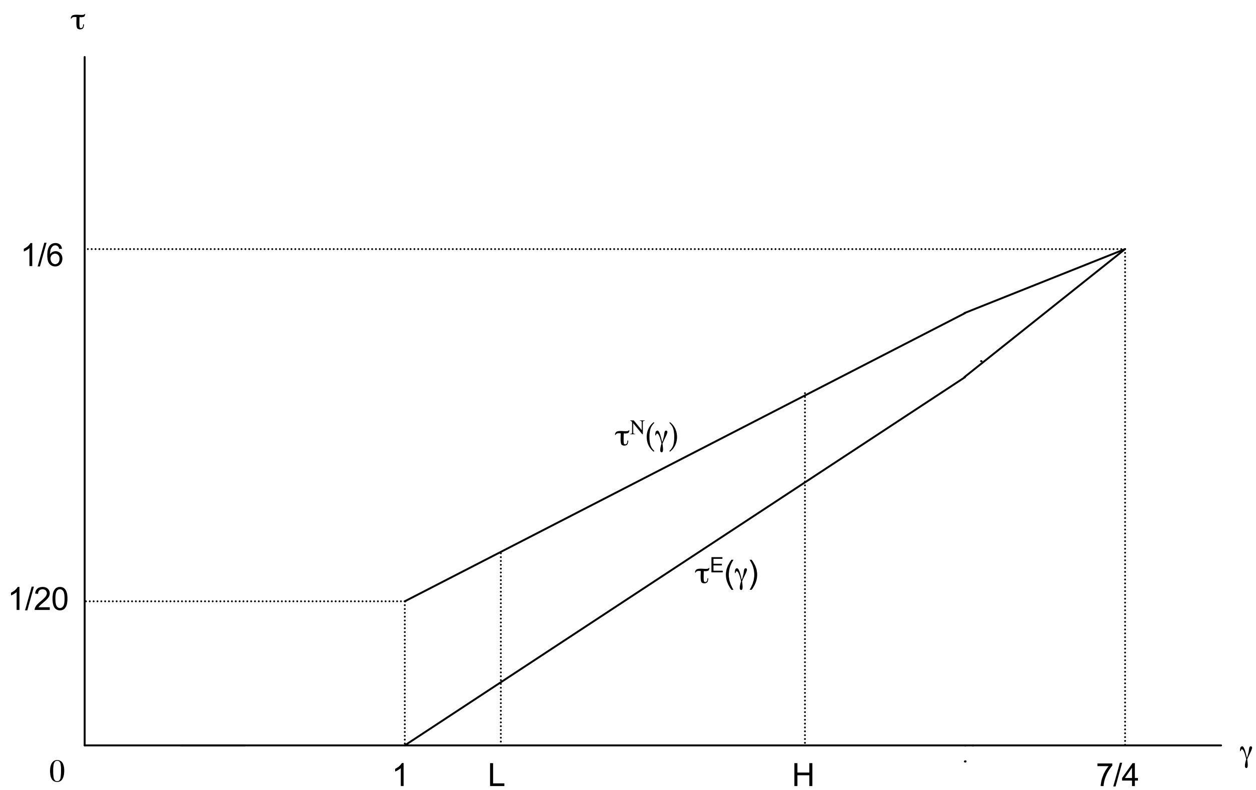

Figure 1
nash and efficient tariffs

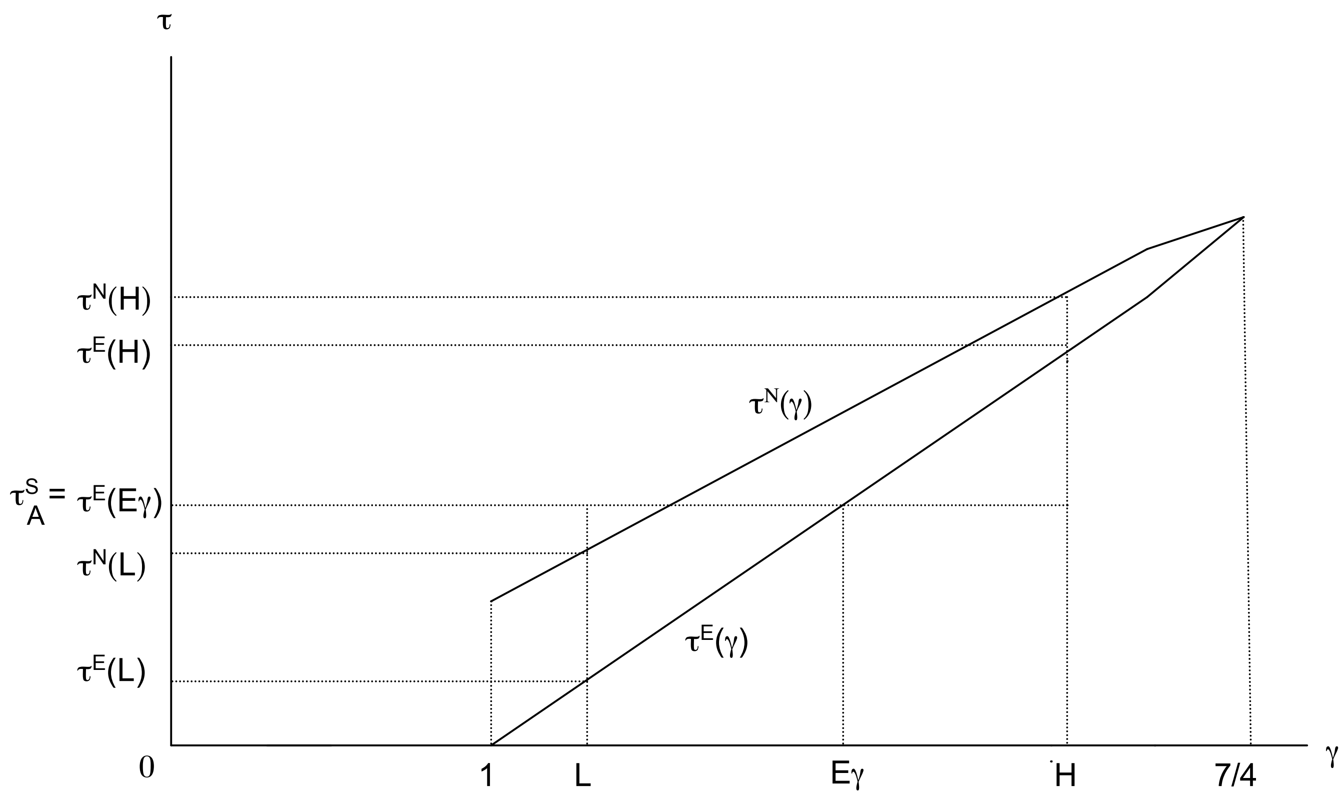

Figure 2
optimal strong binding

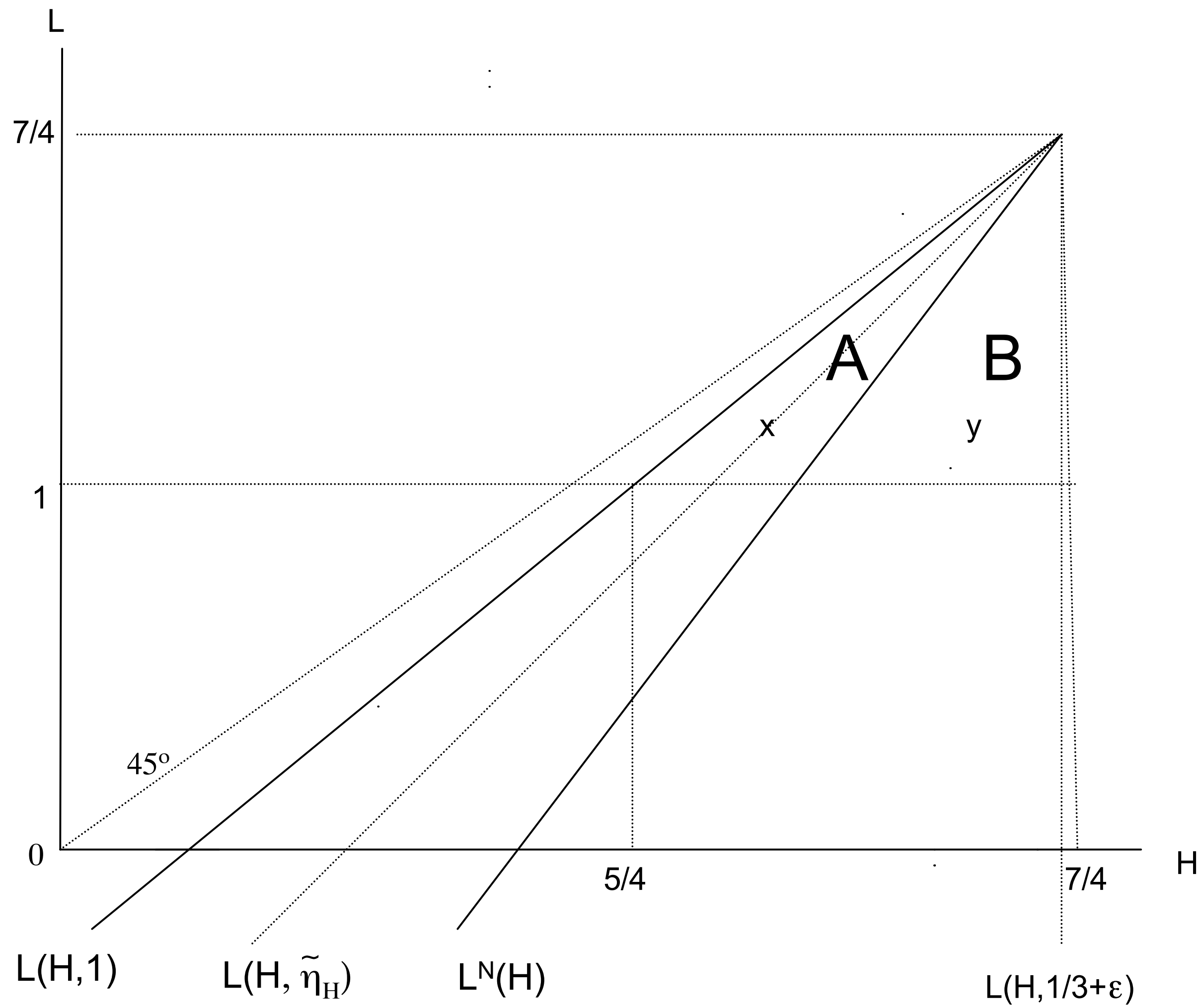

Figure 3
relationship between $L(H,\eta_H)$ and $L^N(H)$

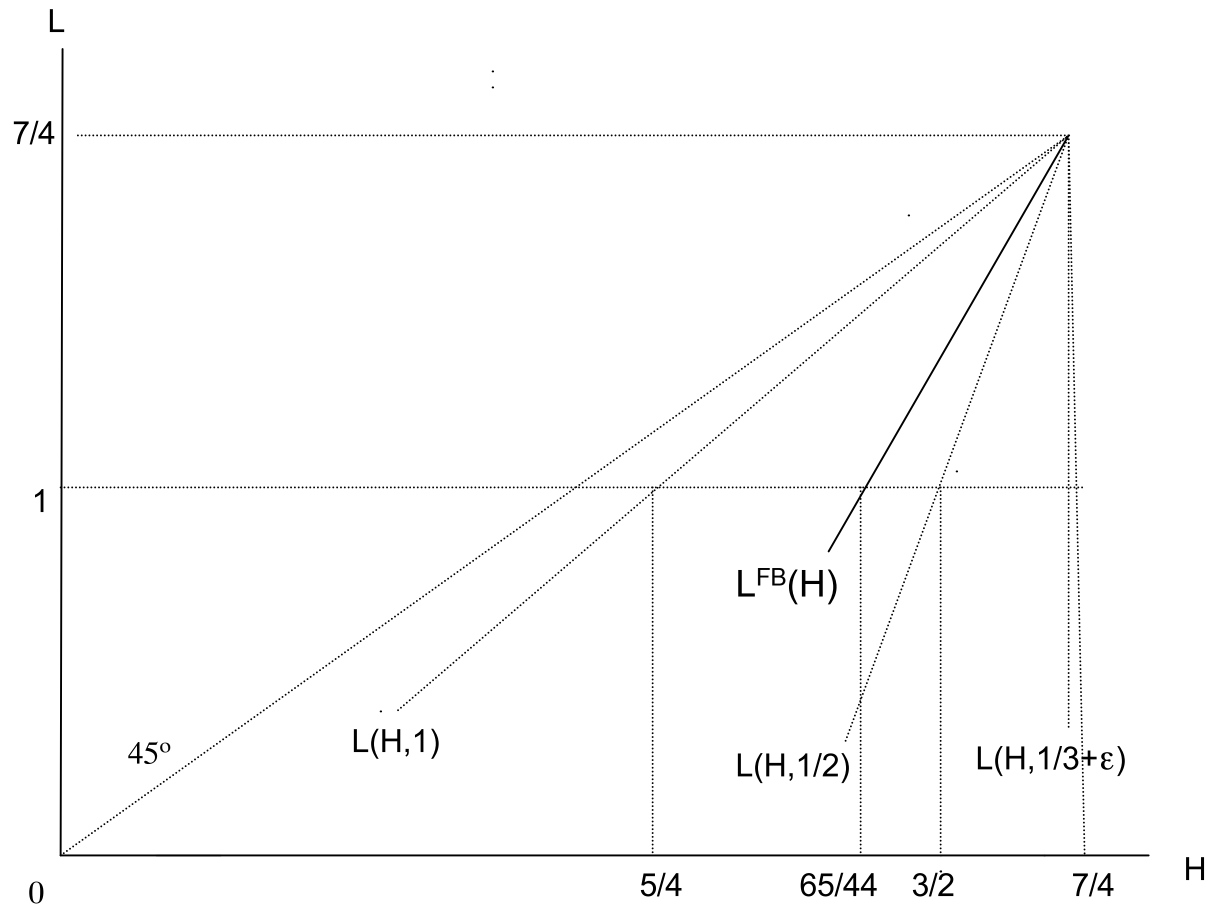

Figure 4
region where efficient tariffs are incentive compatible

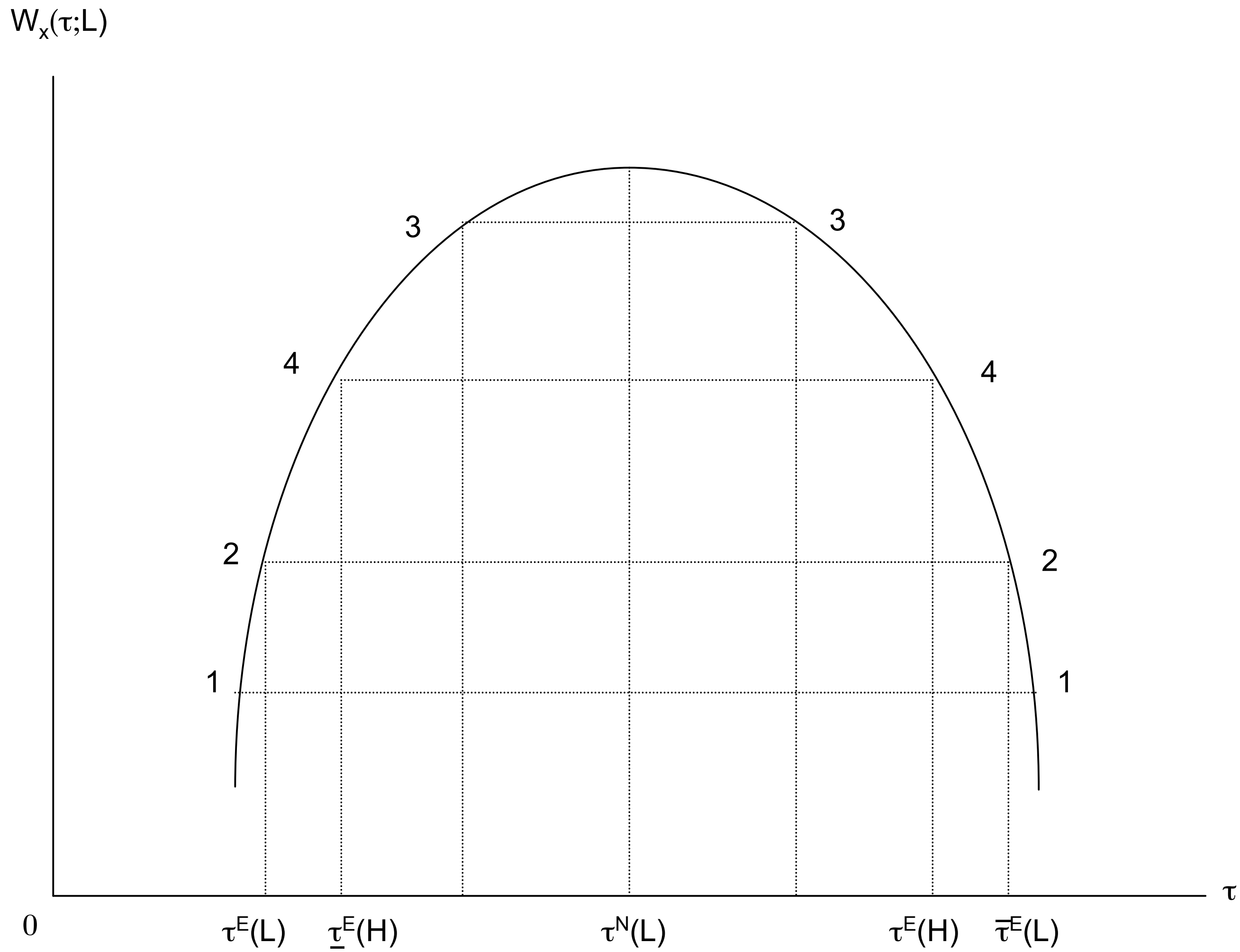

Figure 5
determination of second-best tariffs

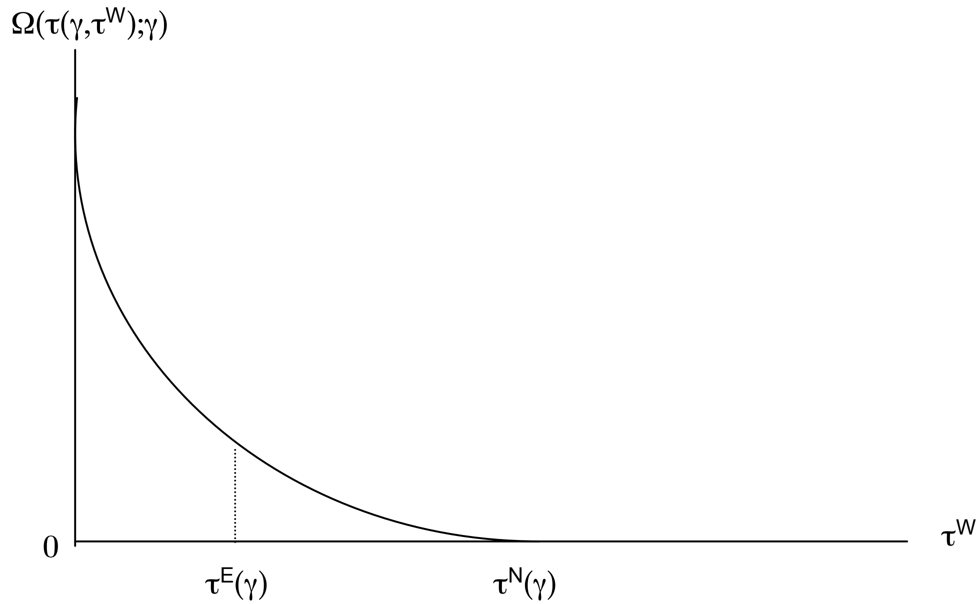

Figure 6a
incentive to cheat

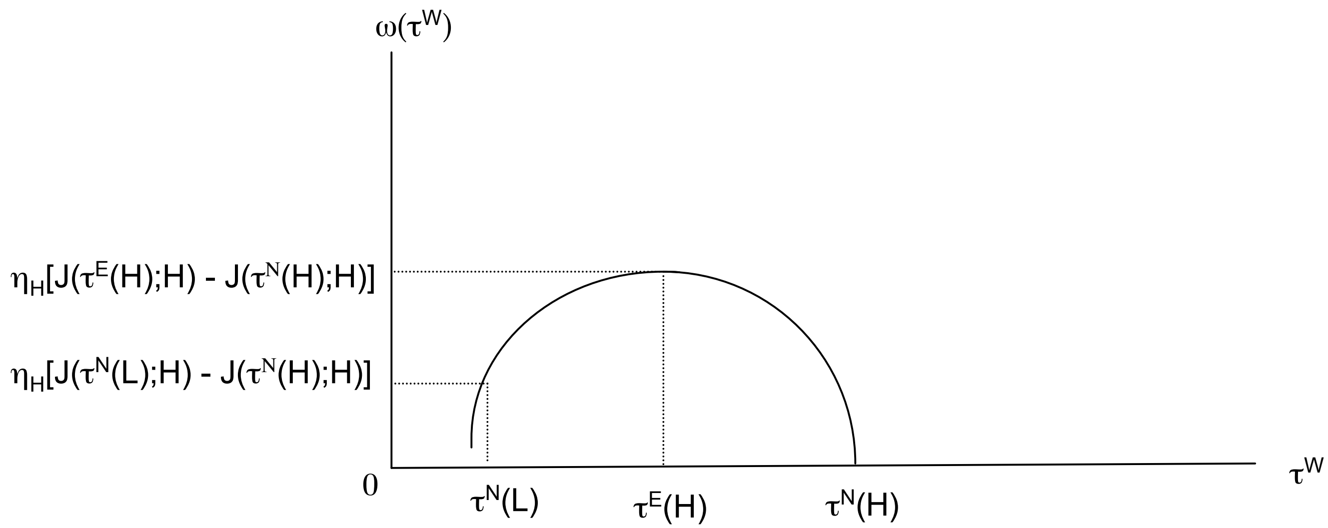

Figure 6b
per-period value of cooperation (when $\omega(\tau^N(L)) > 0$)

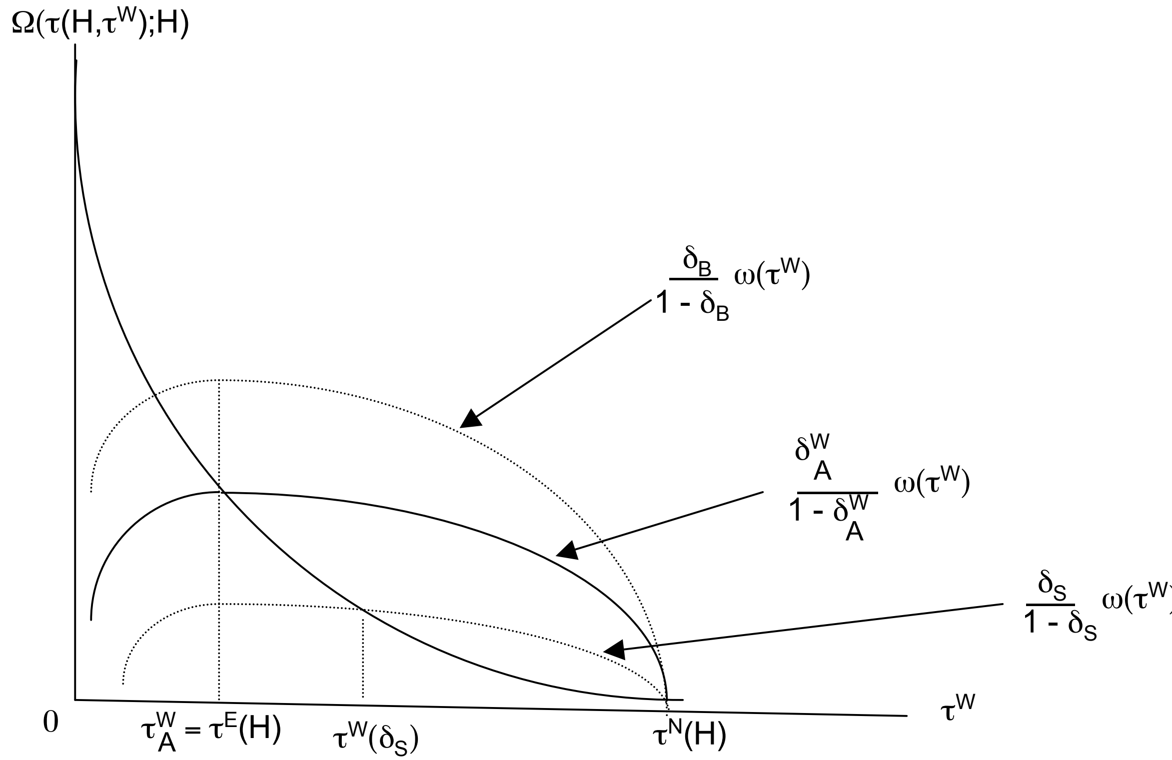

Figure 7
Critical discount factor and most-cooperative tariff (when $\delta_B > \delta^W_A > \delta_S$)

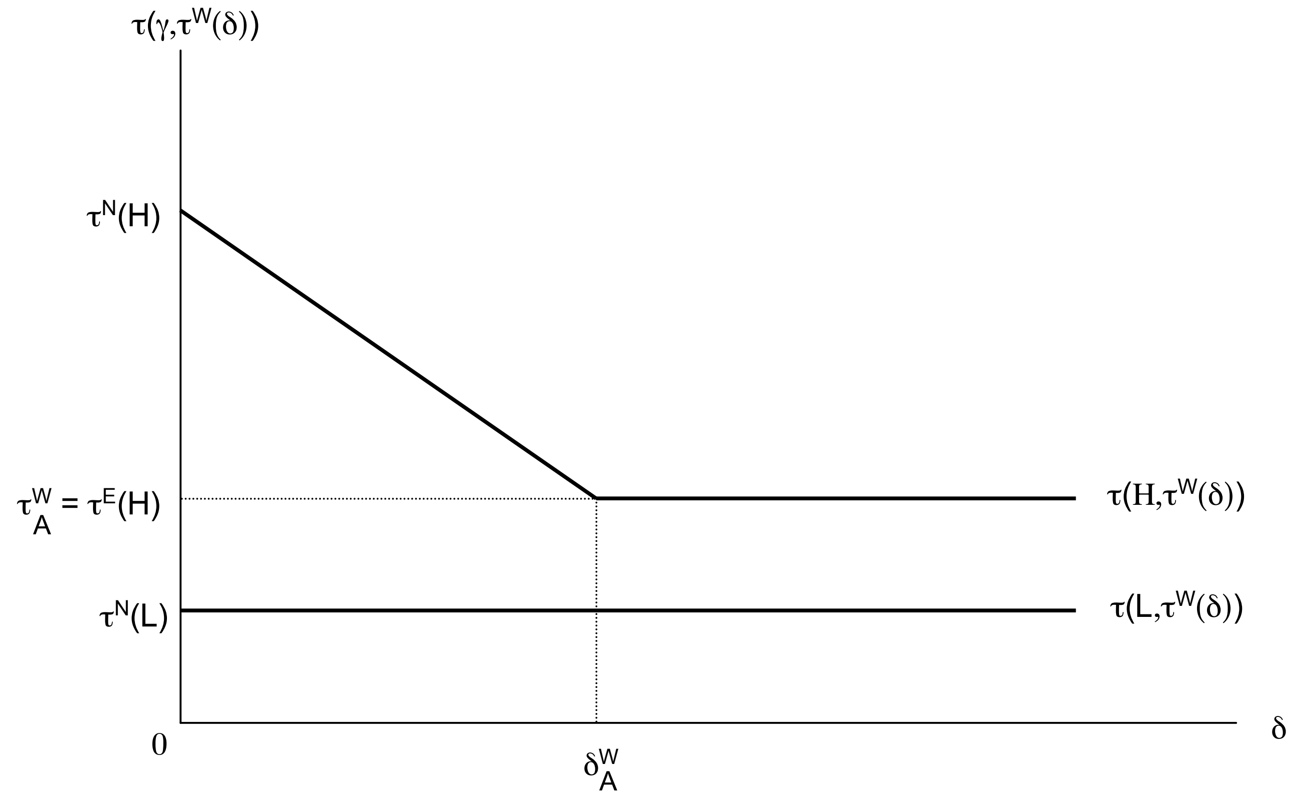

Figure 8
applied tariffs under the most-cooperative weak binding

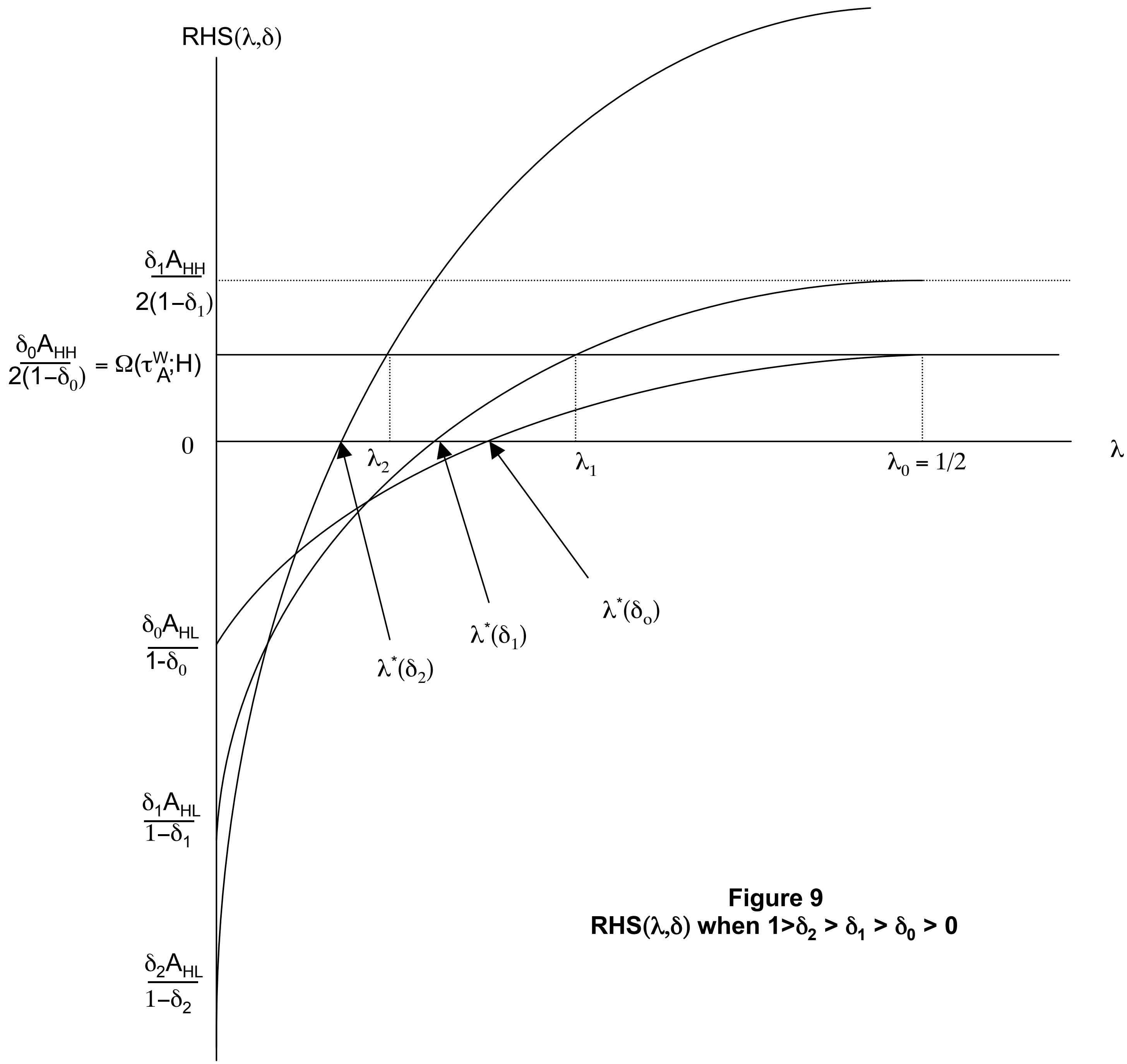

Figure 9
RHS(λ,δ) when $1>\delta_2 > \delta_1 > \delta_0 > 0$