CAN SELF-DESTRUCTIVE TRADE COOPERATION BE OPTIMAL?*

Mathias Herzing†

Abstract

This paper focuses on the impact of hidden information on strategic interaction in the context of trade agreements. In an infinitely repeated tariff setting game between two symmetric countries, informational asymmetry is introduced by letting the weight a government attributes to present vis-à-vis future payoffs be stochastically determined and non-observable to the trading partner. It is shown that when at least some weight will always be given to future payoffs, cooperation can be infinitely sustained if cooperative tariffs are sufficiently close to the Nash tariff level. If tariffs are further reduced, either cooperation breaks down instantly or it can only be sustained as long as governments are sufficiently patient, with the likelihood of breakdown increasing as the cooperative tariff decreases. In the latter case, governments will thus ex ante face a tradeoff between liberalization and sustainability of cooperation. It is shown that it may be optimal to agree on a degree of liberalization associated with a strictly positive ex ante probability of deviation occurring. In that case, cooperation will break down in finite time, and the optimal agreement will thus be self-destructive.

*Helpful comments from Harry Flam, Henrik Horn and Dan Kovenock are appreciated. Thanks also to Christina Lönnblad for editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

†Department of Economics, Stockholm University, August 2006.

1 INTRODUCTION

A commonly raised objection against too far-reaching trade liberalization is that it may increase the incentive for non-cooperative (protectionist) behavior, and that liberalization should therefore be restrained, in order to meet the objective of making liberalization sustainable over time. The implicit assumption behind such an argument is that there exists a trade-off between liberalization and the sustainability of cooperation and that, while a higher degree of liberalization yields a higher expected short-term return, the loss stemming from an increased risk of a breakdown in cooperation is sufficiently large to outweigh the expected short-term gain. Hence, liberalization should be limited by the requirement of sustaining cooperation and avoiding protectionist measures in all contingencies.

The starting point of this paper is the notion that ex ante, when a trade agreement is negotiated, it is typically impossible to know with certainty how large the incentive to deviate will be, once the agreement is in place. There may be various sources of uncertainty, but this paper will examine a situation where two countries are exposed to a random variable, the realization of which is only privately observable. Due to this informational asymmetry, a government must infer the likelihood of its trading partner choosing cooperation from the commonly known distribution of the random variable, and the degree of liberalization agreed upon.

For the sake of analytical tractability, a model is employed where the weight attributed to present vis-à-vis future payoffs is stochastically determined and non-observable to the trading partner. It is shown that in the presence of this hidden-information-type of uncertainty, the scope for liberalization will decrease. Moreover, it is demonstrated how countries may face a tradeoff, when higher degrees of liberalization are associated with decreasing probabilities of cooperation being maintained. It may nevertheless be optimal to agree on a degree of liberalization such that there is a strictly positive likelihood of cooperation breaking down in finite time, since the short-term gain from increasing liberalization may outweigh the long-term loss of cooperation eventually breaking down.

This last observation has interesting implications for the ongoing debate about how far liberalization should be pushed. As is shown, the fact that liberalization is not sustainable in all contingencies may not necessarily mean that liberalization has gone too far, since such an (self-destructive) agreement may be preferable to a safe agreement under which cooperation can be guaranteed forever.

The following section reviews the literature on strategic interaction under various types of uncertainty. Section 3 introduces the model. The scope for liberalization under uncertainty is examined in section 4, and optimality under uncertainty is addressed in section 5. Section 6 concludes.

2 RELATED LITERATURE

The literature on the implementability and sustainability of trade agreements typically examines the strategic interrelationship between two trading countries that can influence world prices through their import tariffs. The countries are in a Prisoner's Dilemma situation, where both would benefit from mutually reducing tariffs but where, from a short-term point of view, each country prefers to apply its best-response tariff vis-à-vis its trading partner. With repeated interaction between the trading partners, it is possible to sustain lower tariffs, however. In standard fashion, by threatening to punish current-period deviations in future periods, the incentive to deviate can be balanced and cooperation be sustained.

The establishment of cooperation hinges on two factors. First, the discount factor must be sufficiently large for the future loss from being punished to outweigh the current gain from deviating. The lower the cooperative tariffs are set, the larger need the discount factor be. Second, there must be a sufficiently high degree of trust between the two parties for cooperation to be established. A country will opt for cooperation only if it attributes a sufficiently high probability to cooperative behavior by its trading partner; believing that the trading partner will deviate makes deviation the preferred choice. The lower the cooperative tariffs are set, the larger must the degree of trust in the trading partner be. Addressing the issue of creating cooperative behavior, Dixit (1987) notes that if each country attributes some positive probability to the trading partner being willing to establish a Nash superior outcome, it then becomes rational for each country to foster such a belief about itself by applying the cooperative tariff in an initial phase of the repeated game.

Introducing some sort of uncertainty into the conditions, under which decisions of complying with or breaching commitments made under a trade agreement are taken not only makes the analysis more complicated, but may also lead to different implications for the prospects of sustaining a cooperative arrangement. Several attempts have been made to incorporate uncertainty into the Prisoner's Dilemma setting of trade agreements. Obviously it is easy to find close correspondences to the literature on collusion under uncertainty. There are, broadly, three categories of uncertainty that have been addressed in the industrial organization literature and, to a lesser extent, also in the literature on trade agreements.

First, strategically interacting parties may be subject to ex ante uncertainty regarding a commonly observed shock which has an impact on the incentive to deviate from a cooperative arrangement. In Bagwell and Staiger (1990), the case of negotiating an agreement in the presence of trade volume fluctutations is analyzed. Periods of high trade volumes are associated with stronger incentives to deviate so as to make terms-of-trade gains and therefore, a cooperative agreement will have to allow for the cooperative tariff to adjust in order to dampen trade volume fluctuations and hence the incentive to deviate. This type of trade management can thus be seen as an attempt by countries to maintain the self-enforcing nature of existing international cooperation.[1]

Second, uncertainty may also emerge because the strategic partner's actions are unobservable.[2] Assuming that commonly observed outcomes are not only influenced by the actions taken, but also by some stochastic variable, it is even ex post not possible to verify or correctly infer the action taken by the opponent in the previous period. A worse-than-expected outcome can thus be the consequence of either deviation on behalf of the opponent, or a bad realization of the stochastic variable. Riezman (1991) models a trade agreement between two countries in the presence of a random component attached to home imports, reflecting shocks to preferences or endowments, and assumes that protection is not perfectly observable.[3] When import trigger strategies are applied, reversionary (high tariff) episodes are triggered by the random variable.

[1] The setting and the results in Bagwell and Staiger (1990) are similar to Rotemberg and Saloner (1986), who model cooperative arrangements between firms that are subject to ex ante unknown, but commonly observable fluctuations in demand.

[2] It is an underlying assumption in all repeated Prisoner's Dilemma models that the action taken by the opponent is not observable in the current period. In the following period, however, the choice taken by the opponent becomes common knowledge, either directly, or indirectly through inference.

[3] The same type of uncertainty also appears in Hungerford (1991) and Kovenock and Thursby (1992), which both focus on the dispute settlement procedures of the General Agreement on Tariffs and Trade.

These Nash reversions are not the result of deviation against the low-tariff agreement, but necessary to provide the incentives for sustaining cooperative low tariff episodes.[4]

Finally, uncertainty can also arise when each party is exposed to a random variable, the realization of which cannot be observed by the opponent. With regard to this hidden-information type of uncertainty in the context of trade cooperation, the case of one-stage games with implicitly assumed cooperative behavior has been examined by Feenstra (1987) and Feenstra and Lewis (1991). Given the underlying assumption that countries prefer to cooperate and apply agreed-upon policies, the problem addressed is how hidden information creates incentives to misrepresent in order to make gains. While these two contributions focus on the effect of hidden information on outcomes in cooperative games, Jensen and Thursby (1990) investigate the effect of private information on noncooperative equilibria and how the incentives for governments to establish tariff reputations might be influenced.

Only recently has the impact of hidden information on cooperation in repeated non-cooperative games been explored. Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004) elaborate on the impact of hidden information in infinitely repeated Bertrand games between competing firms when prices are publicly observed, but cost shocks are only privately observed. Martin and Vergote (2005) apply a repeated setting with unobserved political preferences, in which antidumping is used to let tariffs adjust to changes in these preferences. They show that truthfulness about political preferences can be achieved by letting present actions influence expected future payoffs through retaliation, and that the retaliatory use of antidumping may then improve welfare if static rules governing its use are adopted. These results suggest that when the use of some instruments are restricted, the strategic or retaliatory use of the remaining ones, such as antidumping, may be the most efficient way of dealing with hidden information.

While the present paper also attempts to shed light on the impact of hidden information on cooperative outcomes in an infinitely repeated non-cooperative setting, its main focus is the trade-off between the degree and the sustainability of liberalization. The starting point is an

[4] This result corresponds to that obtained in Green and Porter (1984), in which low realizations of the price trigger reversionary episodes when two firms strategically interact in a Cournot duopoly that is affected by unobservable demand shocks influencing the price level. See also Abreu, Pearce and Stacchetti (1986), who modify and generalize the Green-Porter model.

agreement only specifying cooperative tariff levels and not allowing for any kind of direct or indirect transfers, neither within nor across periods. It is also implicitly assumed that no other trade-distorting instrument than tariffs is available. Hence, the agreement does not include any safeguard provisions, like the antidumping instrument in Martin and Vergote (2005).[5] A government facing strong incentives to deviate will thus have no other choice than raising tariffs, thereby breaching the commitment to liberalization. Naturally, this is a major simplification, but it is made to gain insights into what degree of liberalization will be chosen ex ante, when ex post gains from deviation may threaten the sustainability of liberalization.

Hidden-information-type of uncertainty in the context of trade cooperation can be described as a situation where a government is exposed to some random variable that only it, and no one else, can observe, such that the true incentives faced by the government with regard to choosing to comply with or deviate from liberalization commitments are unknown to the trading partner(s). There exist various ways of introducing this type of uncertainty in the government's objective function. Baldwin (1987) introduces a politically realistic objective function (PROF) and shows it to be equivalent to the payoff functions derived from a wide range of political economy models. The PROF attributes different weights to consumer surplus, tariff revenues and different industry profits. A commonly used way of introducing uncertainty is to let one of these weights, typically profits of an import-competing sector, be randomly determined.

A simpler and analytically more tractable way of investigating the impact of uncertainty on the conditions for strategic interaction and cooperation in an infinitely repeated setting is to let the government weigh current-period and future-period payoffs and let one of these weights (and hence, the relative weight) be randomly determined. In fact, attributing more weight to profits in the import-competing sector is similar to giving more weight to present vis-à-vis future payoffs. In both cases the incentives for protection increase.

Infinitely repeated games with stochastically determined discount factors have previously been examined by Baye and Jansen (1996).[6] In their model both parties are exposed to the same discount factor in every

[5] The impact of introducing a safeguard, whereby deviations from cooperation are permitted, in an infinitely repeated tariff setting game is addressed in Herzing (2005).

[6] Bernheim and Dasgupta (1995) analyze the impact on cooperation in infinitely repeated games with asymptotically falling discount factors.

period, unlike in the present model where the two governments' realizations of the weight attributed to present payoffs are uncorrelated and only privately observable. Baye and Kovenock (2004) model a Bertrand duopoly supergame where each firm's per-period discount factor is stochastically determined and only privately observable at the time prices are set. Their setting differs from the present one in two respects. First, the discount factor can only take on two values, while in the presenet model the weight attributed to present payoffs is drawn from a continuum. Second, the nature of strategic interaction is different in the present model, because tariff setting is strategically similar to Cournot competition. Under Bertrand competition, cooperation is a matter of applying the monopoly price or the Nash price, while Cournot competition allows for a continuum of cooperative quantities to choose among. Thus, the conclusions drawn from this model can be applied to the case of a Cournot duopoly.

3 THE MODEL

Let there be two symmetric countries, home and foreign (distinguished by an asterix), each with one sector. In standard fashion the two countries interact in an infinitely repeated tariff setting game, and each government is assumed to be subject to a random variable that is only privately observable and i.i.d. across countries and periods. Due to symmetry, it suffices to consider the home country. The random variable is assumed to enter, such that the home government's payoff function W^{δ} in the current period is given by:

$$W^{\delta} = \delta w + (1 - \delta)v,$$

where w is the payoff of the current period, v the expected discounted future flow of payoffs, and $\delta \in [\delta_{\min}, \delta_{\max}] \subseteq [0, 1]$ the weight attributed by the government to present payoffs, the residual $1 - \delta$ being the weight given to future payoffs. Let $\overline{\delta} \equiv E(\delta) \in [\delta_{\min}, \delta_{\max}]$ be the expected value of δ. Ex ante, there is uncertainty concerning what weight a government will attribute to present payoffs in every period. The density function $\varphi(\delta)$, which is assumed to be continuous for $\delta \in [\delta_{\min}, \delta_{\max}]$, and the associated cumulative distribution function $\Phi(\delta)$ are common knowledge, however. Hence, $\overline{\delta}$ can be seen as a measure for the shortsightedness of a government (a high value of $\overline{\delta}$ implies a low discount factor), while $\delta_{\max}$ represents the most myopic realization of δ possible.

For analytical tractability, a partial equilibrium setting where the current-period payoff w is additively separable in the home tariff t and the foreign tariff t^*, is assumed

$$w(t, t^*) = u(t) + \widehat{u}(t^*).$$

There exists a best-response function $t_D(t^*) \equiv \arg\max_t w(t, t^*)$. From the additive separability of w, it immediately follows that the within-period best-reply tariff t_D is independent of t^* and hence constant.

In the absence of cooperation, both countries apply the optimal tariff vis-à-vis each other, i.e. $t = t^* = t_D$, and both receive the Nash equilibrium current-period payoff $w_N = w(t_D, t_D)$. Since the current-period payoff in every future period will be equal to w_N, the government's payoff in the absence of cooperation will be given by $W_N^\delta = w_N$, i.e. it will be independent of δ.

However, both countries can be made better off by agreeing to mutually lower their tariffs. What prevents the implementation of a cooperative tariff $t_C < t_D$, associated with a current-period payoff $w_C = w(t_C, t_C) > w_N$, is that the two countries are stuck in a Prisoner's Dilemma. If one country decides to break its commitment by applying the optimal tariff vis-à-vis its trading partner, it gets the current-period payoff $w_D = w(t_D, t_C) > w_C$, while its trading partner receives the sucker's payoff $w_S = w(t_C, t_D) < w_N$. Thus, both countries will apply t_D, and their current-period payoff will be given by w_N in the one-shot game.

It immediately follows that w_C has a unique maximum for $t_C = t_C^{opt} < t_D$ and that there exists a $t' < t_C^{opt}$ such that $w_C = w_N$ for $t_C = t'$. Thus, $w_C > w_N$ if and only if $t' < t_C < t_D$. It is also straightforward that the deviator's payoff w_D decreases unambiguously in t_C. The decrease in w_D is equal to the decrease in w_C (i.e. w_D and w_C are tangent) at $t_C = t_D$ and unambiguously larger for $t_C < t_D$. Thus, $w_D - w_C$ decreases, and it does so at a decreasing rate as t_C increases.

Define

$$\tau \equiv \frac{w_D - w_C}{w_D - w_N}.$$

It is easily shown that $\lim_{t_C \to t_D} \tau = 0$ and that τ increases monotonously as t_C decreases below t_D. Hence, τ can be seen as a measure of trade liberalization. A low value of t_C corresponds to a high value of τ and thus, a high degree of trade liberalization. Since $w_C^{opt} > w_N$, the optimal degree of liberalization τ_{opt}, corresponding to $t_C = t_C^{opt}$, is strictly smaller than unity and, because $w_C = w_N$ for $t_C = t'$, the degree of liberalization τ', corresponding to $t_C = t'$, equals unity. The relevant range of cooperative tariffs to consider is given by (t', t_D), corresponding to degrees of liberalization in the range $(0, 1)$. In this range, it is the case that $w_D > w_C > w_N > w_S$, and strategic interaction is thus of Prisoner's Dilemma type.

As mentioned above, mutual deviation is the only equilibrium outcome in the one-shot Prisoner's Dilemma game. Under an infinite horizon, it is possible to create cooperation, however.[7] By threatening to punish deviations and thus associate the one-period gain from deviating with a future loss, it is possible to induce cooperative behavior. There exist many different ways of conceiving punishment phases. For simplicity grim-trigger strategies will be considered throughout the paper. Hence it will be assumed that deviations from the cooperative tariff will trigger infinite reversion to the Nash equilibrium.

In the absence of uncertainty, $\delta = \overline{\delta}$ and, assuming a propensity for cooperative behavior across countries[8], choosing cooperation yields $W_C = \overline{\delta} w_C + (1 - \overline{\delta}) w_C = w_C$, while opting for deviation yields $W_D = \overline{\delta} w_D + (1 - \overline{\delta}) w_N$. Cooperation is thus sustainable, if and only if

$$w_D - w_C \leq \frac{1 - \overline{\delta}}{\overline{\delta}} (w_C - w_N). \tag{1}$$

The left-hand side represents the short-term (current-period) gain from deviation, while the right-hand side represents the expected long-term loss from deviation. Since $w_C > w_N$ under an agreement, rearranging terms yields the following relationship between the degree of liberalization and the discount factor, which is given by $1 - \overline{\delta}$, i.e. the weight attributed to the future flow of payoffs.

$$(1) \Leftrightarrow \frac{w_D - w_C}{w_C - w_N} \leq \frac{1 - \overline{\delta}}{\overline{\delta}} \Leftrightarrow \frac{\tau}{1 - \tau} \leq \frac{1 - \overline{\delta}}{\overline{\delta}} \Leftrightarrow \tau \leq 1 - \overline{\delta} \equiv \tau_{\max} \tag{1'}$$

Equation (1') tells us that in order to sustain cooperation, t_C can only be reduced to the degree where τ does not exceed the discount factor $1 - \overline{\delta}$. A lower $\overline{\delta}$, i.e. a higher discount factor, implies that the

[7] In fact, cooperative behavior can be established in all periods but the last in a finitely repeated game, if there exist multiple Nash equilibria. As pointed out by Dixit (1987), this is the case when a tariff setting game includes the possibility of reversion to autarky. However, it is not true if the two countries interact strategically by setting quotas. In that case, an infinite horizon is necessary for cooperation to be possible, because autarky is the only static Nash equilibrium (see Tower (1975)).

[8] It is easily established that both countries choosing to deviate constitutes an equilibrium outcome, independent of the degree of liberalization and the weight attributed to current-period payoffs.

upper bound for liberalization $\tau_{\max}$ increases, and it is thus possible to sustain a lower t_C. The restriction given by (1') is incorporated in the negotiations concerning the cooperative tariff level and thus, imposes an upper limit on the scope for liberalization.

Two well-known results immediately follow from condition (1'). First, it is always possible to find some $\tau > 0$ ($t_C < t_N$) that is sustainable for $\overline{\delta} < 1$ (i.e. a strictly positive discount factor). Second, the optimal degree of liberalization τ_{opt} can be sustained, if governments are sufficiently patient, i.e. if the weight attributed to current payoffs is not too large ($\overline{\delta} \leq 1 - \tau_{opt}$).

4 THE SCOPE FOR TRADE LIBERALIZATION UNDER UNCERTAINTY

Introducing uncertainty about the weight the trading partner attributes to current vis-à-vis future payoffs significantly complicates the analysis. The incentive to deviate will not only depend on the ex ante unknown realization of δ and the degree of liberalization, but also on the likelihood p of the trading partner choosing cooperation. In equilibrium, a government's ex ante probability of opting for cooperation must equal its belief regarding the other government's likelihood of choosing cooperation, i.e. beliefs must be consistent.

This also applies in the absence of uncertainty. However, consistent solutions are much easier to derive under certainty. Any prior regarding the likelihood of cooperative behavior of the trading partner either results in deviation or, possibly (i.e. for sufficiently low degrees of liberalization), cooperation being a consistent solution.[9] Hence, cooperation over the infinite horizon is only a matter of establishing trust, such that both countries coordinate on the cooperative solution.

In the face of uncertainty, it is also the case that a propensity for maximally cooperative behavior is required for the most cooperative outcome to be attained, but it may not be sufficient to sustain cooperation in all contingencies. Knowing that the trading partner may deviate under certain circumstances feeds back into the decision for when opting for cooperation is optimal. Anticipating that the foreign country's government must similarly infer its optimal strategy from a belief concerning the likelihood of cooperation by the home country's government,

[9] Actually, whenever cooperation is a consistent outcome, there also exists a mixed strategy (i.e. randomizing between cooperation and deviation) that is consistent.

a process of updating of initial priors will yield consistent solutions. The derivation of consistent solutions turns out to be analytically non-trivial.

As above in the case of certainty, it will be assumed that grim-trigger strategies are applied. Let v_C and v_D be the continuation values of the game if both countries cooperate and if at least one country deviates, respectively. For any strictly positive degree of liberalization, the gain from deviating, denoted by Γ, is thus given by

$$\Gamma = p[\delta(w_D - w_C) + (1-\delta)(v_D - v_C)] + (1-p)\delta(w_N - w_S).$$

Since w is additively separable, the one-period gain of deviating is independent of what action the trading partner takes, and in particular $w_N - w_S = w_D - w_C$. Therefore, any decision will solely depend on the domestic shock and the likelihood attributed to cooperation being chosen by the trading partner. Hence, the gain from opting for deviation is expressed as follows

$$\Gamma = \delta(w_D - w_C) + p(1-\delta)(v_D - v_C).$$

The condition for when cooperation is chosen is thus given by

$$\Gamma \leq 0 \Leftrightarrow \delta \leq \frac{p(v_C - v_D)}{w_D - w_C + p(v_C - v_D)} \equiv \eta.$$

A country will opt for cooperation as long as realizations of δ are smaller than the threshold value η. This threshold value, in turn, implies an ex ante likelihood of this country choosing cooperation of $\Phi(\eta)$. In fact, η can be regarded as a reaction function of p, i.e. the probability attributed to the trading partner choosing cooperation. By introducing the concept of consistency in beliefs regarding the likelihood of cooperation being chosen, solutions for η can be derived by treating p as a prior regarding the likelihood of the trading partner opting for cooperation and updating it.

Symmetry across countries and consistency require that the probability of the trading partner choosing cooperation must equal the implied likelihood of the own country opting for cooperation, i.e. $p = prob(\delta \leq \eta) = \Phi(\eta)$. Consistent solutions are thus given by solutions to the following equation

$$\eta = \frac{(v_C - v_D)\Phi(\eta)}{w_D - w_C + (v_C - v_D)\Phi(\eta)}. \tag{2}$$

The continuation values v_C and v_D are given by (see the Appendix for derivation)

$$v_D = w_N \tag{3}$$

$$v_C(\tau,\eta) = w_N + \frac{\bar{\delta}\Phi(\eta) - \tau\int_{\delta_{\min}}^{\eta} \delta d\Phi}{1-\Phi(\eta)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi]}(w_D - w_N). \tag{4}$$

Hence, v_C is affected by both the degree of liberalization and the associated threshold value, above which deviation will be chosen. Plugging the expressions for v_C and v_D into equation (2) and rearrangement of terms then yield the following equation, henceforth referred to as the consistent solution equation (CSE)

$$\eta = \Phi(\eta)\frac{\bar{\delta}\Phi(\eta) - \tau\int_{\delta_{\min}}^{\eta} \delta d\Phi}{\bar{\delta}\Phi^2(\eta) + \tau[1-\Phi^2(\eta)]} \equiv f(\tau,\eta). \tag{5}$$

The right-hand side of the CSE is given by

$$f(\tau,\eta) = \begin{cases} 0 & \eta \leq \delta_{\min} \\ \Phi(\eta)\frac{\bar{\delta}\Phi(\eta)-\tau\int_{\delta_{\min}}^{\eta}\delta d\Phi}{\bar{\delta}\Phi^2(\eta)+\tau[1-\Phi^2(\eta)]} & \eta \in (\delta_{\min}, \delta_{\max}) \\ 1-\tau & \eta \geq \delta_{\max} \end{cases}.$$

Define $\Omega(\tau) \equiv \{\eta | \eta = f(\tau,\eta)\}$ as the set of consistent solutions. It is easy to see that $\Omega(\tau) \neq \oslash$ for any degree of cooperation. Since $f(0) = 0$, $\eta_0 \equiv 0$, implying that cooperation is never chosen, is a consistent solution for any degree of liberalization irrespective of the realization of δ. Naturally, this is also true in the case of no uncertainty. It is important to emphasize that the sustainability of cooperation in infinitely repeated Prisoner's Dilemma games does not only depend on the discount factor, but also on the prior regarding the likelihood of the opponent opting for cooperation. By updating the prior, a consistent best response is derived. Hence, the degree of trust in the opponent is of crucial importance for sustaining a cooperative regime. Having low faith in the opponent results in deviation being the best response, under certainty as well as under uncertainty.[10]. The following lemma follows by introspection of the CSE.

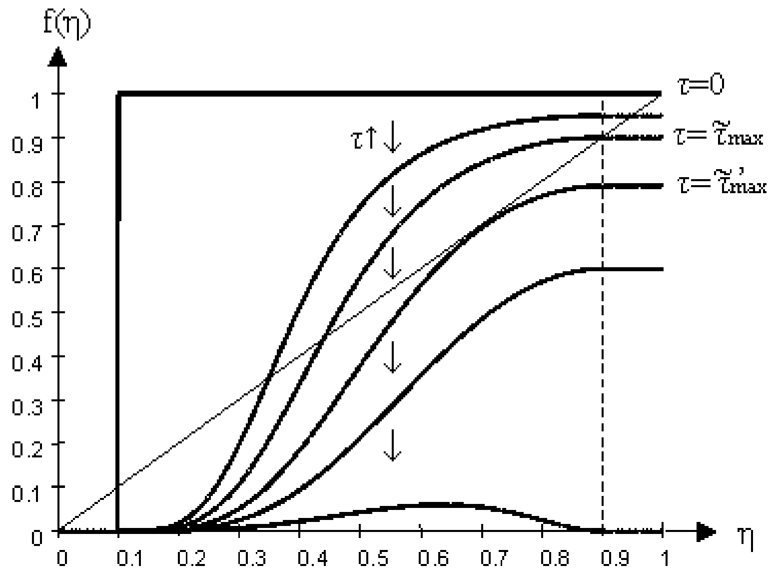
[10] The degree of trust initially required, i.e. the value of the prior before updating, may naturally differ under certainty and uncertainty.

Lemma 1 *For sufficiently low degrees of liberalization ($\tau \leq 1 - \delta_{\max} \equiv \widetilde{\tau}_{\max}$), there exists a consistent solution ($\eta = 1 - \tau \geq \delta_{\max}$) associated with a likelihood of cooperation equal to unity. The range of degrees of liberalization, for which cooperation is always chosen, is, however, unambiguously smaller under uncertainty than under certainty ($\widetilde{\tau}_{\max} < \tau_{\max}$).*

Hence, it is also possible to sustain cooperation over the infinte horizon for sufficiently low degrees of liberalization, just as in the case of uncertainty, if $\delta_{\max} < 1$. However, the range of degrees of liberalization, for which this is true, is smaller in the case of uncertainty, because $\overline{\delta} < \delta_{\max}$.

If the degree of liberalization is pushed further than $\widetilde{\tau}_{\max}$, always opting for cooperation can no longer be a consistent solution, because the incentive to deviate will be too strong for large realizations of δ. There are two possibilities. Either cooperation completely breaks down, i.e. deviation is the preferred option for any δ, as in the case of certainty for $\tau > \tau_{\max}$; or it will be the case that cooperation is only chosen for sufficiently low realizations of δ, i.e. the threshold value η will lie in the interval $(\delta_{\min}, \delta_{\max})$, implying a probability of cooperation being chosen strictly smaller than one, but also strictly larger than zero. In the latter case, deviation will thus occur in finite time. Let $\widetilde{\tau}'_{\max}$ be the highest degree of liberalization for which the ex ante probability of choosing cooperation is strictly positive.

In what follows, agreements under which cooperation can be sustained forever will be referred to as safe, while agreements under which the ex ante likelihood of cooperation being chosen is strictly positive, but also strictly smaller than one, will be referred to as self-destructive. Whether there exist self-destructive agreements crucially depends on how the CSE is affected by pushing τ above $\widetilde{\tau}_{\max}$. Continuity of the density function implies that $f(\eta)$ is continuous in η for $\tau > 0$. The figure below demonstrates the impact of τ on the solution(s) of the CSE. Using a symmetric density function and setting $\overline{\delta} = 0.5$, $\delta_{\min} = 0.1$ and $\delta_{\max} = 0.9$, $f(\eta)$ is plotted for different values of τ (thick lines). Consistent solutions, i.e. solutions to the CSE, are given by intersections with the upward-sloping thin line, which represents the left-hand side of the CSE. The dashed vertical line denotes $\delta_{\max}$; any intersections to its right thus imply a solution with an associated probability of cooperation being chosen equal to one.



When the degree of liberalization is zero[11], $f(\eta) = 0$ for $\eta \in [0, \delta_{\min}]$ and $f(\eta) = 1$ for all $\eta \in (\delta_{\min}, 1]$. Hence, $\eta_0 = 0$ and $\eta = 1$ are consistent solutions. As can be easily inferred from the CSE, letting τ increase above zero will lead to a decrease in $f(\eta)$ for all $\eta \in (\delta_{\min}, 1]$. Hence, $f(\eta) < 1$, but there exist $\eta \in (0, 1)$, for which $f(\eta) > \eta$. Thus, there will be at least two intersections between η and $f(\eta)$ in this interval. Let $\eta_1 \equiv \max\{\eta \in [0, 1] | \eta = f(\eta)\}$ be the largest and $\eta_2 \equiv \min\{\eta \in (0, 1] | \eta = f(\eta)\}$ the smallest strictly positive solutions to the CSE, respectively. An increase in τ will reduce $f(\eta)$ for all $\eta \in (\delta_{\min}, 1]$ and thus, lead to a decrease in η_1 and an increase in η_2. As long as $\tau \leq \widetilde{\tau}_{\max}$, $\eta_1 \geq \delta_{\max}$ and always cooperating is a consistent solution (as previously stated in Lemma 1). As τ is increased beyond $\widetilde{\tau}_{\max}$, η_1 falls below $\delta_{\max}$, while η_2 continues to increase.[12] Eventually, η_1 and η_2 will coincide, i.e. $f(\eta) < \eta$ for all $\eta > 0$ except $\eta = \eta_1 = \eta_2$ and $f(\eta)$ will be tangent to η for $\eta = \eta_1$. Naturally the degree of liberalization at which η_1 and η_2 coincide is $\widetilde{\tau}'_{\max}$, the highest degree of liberalization, for which the ex ante probability of choosing cooperation is strictly positive. Pushing the degree of liberalization beyond $\widetilde{\tau}'_{\max}$ will lead to $f(\eta) < \eta$ for all $\eta > 0$, i.e. $\eta_0 = 0$ will be the only consistent solution. Hence, η_2 no longer exists and η_1 coincides with η_0 for $\tau > \widetilde{\tau}'_{\max}$.

[11]Note that the plot for $\tau = 0$ is actually the one obtained for $\lim_{\tau \to 0} f(\eta)$.

[12]It may also be the case that $\eta_1 = 0$ and η_2 is non-existent for any $\tau > \widetilde{\tau}_{\max}$. Which case applies will be discussed in detail later.

Generally, it will thus be the case that for sufficiently low degrees of liberalization, there exist (at least) three different solutions, η_0, η_1 and η_2. Intuitively, the largest solution to the CSE should be the preferred choice, because it yields the highest ex ante likelihood of cooperation being chosen. The following lemma shows that this is indeed the case.

Lemma 2 *Let η_1 and η_2 be consitent solutions. If $\eta_1 > \eta_2$, then $v_C(\tau, \eta_1) > v_C(\tau, \eta_2)$.*

Proof. See the Appendix. ■

Henceforth, it will be assumed that governments wish to behave as cooperatively as possible and thus apply η_1 for any given degree of liberalization, because this yields the largest continuation value among all consistent solutions. Thus, each government has an interest in fostering a belief about itself acting as cooperatively as possible, such that η_1 can be derived.[13]

Naturally, there are many possibilities for the shape of the f-function. But to derive a sufficient condition for when self-destructive agreements are feasible, it suffices to examine the case when there exist at most three solutions to the CSE, as in figure 2. Whether it is possible to increase liberalization beyond $\widetilde{\tau}_{\max}$ without cooperation breaking down, will in this case depend on the marginal likelihood of the most myopic realization of δ occurring, as demonstrated by the following proposition.

Proposition 1 *Self-destructive agreements exist (i.e. $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$), if*

$$\varphi(\delta_{\max}) < \frac{1}{(1-\delta_{\max})(1+\frac{\delta_{\max}}{\overline{\delta}})} \equiv \overline{\varphi(\delta_{\max})}. \tag{6}$$

Proof. See the Appendix. ■

The intuition behind this result is the following. A low density for $\delta = \delta_{\max}$ implies that a threshold value η slightly smaller than $\delta_{\max}$ will be associated with a probability of cooperation being chosen negligibly smaller than unity. Hence, the effect on the likelihood of cooperation is only negligibly different from when there is no uncertainty, in which case cooperation can be sustained when τ is marginally increased beyond $\widetilde{\tau}_{\max}$, because $\tau_{\max} > \widetilde{\tau}_{\max}$. However, if $\varphi(\delta_{\max})$ is too large, this will no longer be the case. The probability of cooperation being chosen will fall

[13] The issue of how wide a range of priors support any of the solutions of the CSE and hence, how sensitive each solution is to changes in the prior, is not addressed here.

by sufficiently much to make opting for deviation the only consistent solution.[14]

It is straightforward that a similar condition as stated in the previous proposition, can easily be derived for the case when there exist more than three solutions for the CSE when $\tau = \widetilde{\tau}_{\max}$. For self-destructive agreements to be feasible as τ is increased beyond $\widetilde{\tau}_{\max}$, it is necessary that there exists at least one consistent solution $\eta > \delta_{\min}$, for which the density is sufficiently low.

From condition (6) we can infer that $\overline{\varphi(\delta_{\max})}$ increases unambiguously in $\overline{\delta}$ and increases in $\delta_{\max}$ for $\delta_{\max} \geq \frac{1-\overline{\delta}}{2}$. Thus, the scope for self-destructive agreements increases as $\overline{\delta}$ and/or $\delta_{\max}$ attain large values. Letting $\delta_{\max}$ go to one, $\overline{\varphi(\delta_{\max})}$ becomes infinitely large, which implies that self-destructive agreements will always exist. Hence, when there is a positive marginal probability that governments can become entirely short-sighted, self-destructive agreements are always feasible. In fact, in this case all agreements are self-destructive for any strictly positive degree of liberalization.

5 THE OPTIMAL DEGREE OF LIBERALIZATION UNDER UNCERTAINTY

In the previous section, a sufficient condition for when liberalization can be pushed beyond $\widetilde{\tau}_{\max}$ without cooperation instantly breaking down was derived This section explores what degree of liberalization is optimal. First, consider the case when the optimal degree of liberalization under certainty τ_{opt} can be attained with an associated likelihood of cooperation being chosen equal to unity (i.e. $\widetilde{\tau}_{\max} \geq \tau_{opt}$). In this case it is obviously optimal to apply τ_{opt}, since increasing τ beyond τ_{opt} will lead to a fall in v_C.

Proposition 2 *If $\widetilde{\tau}_{\max} \geq \tau_{opt}$, then the optimal degree of liberalization under uncertainty is equal to the optimal degree of liberalization under certainty ($\widetilde{\tau}_{opt} = \tau_{opt}$).*

Proof. First, consider $\tau \leq \widetilde{\tau}_{\max}$. In this case $\eta(\tau) \geq \delta_{\max}$ and hence, $v_C(\tau, \eta(\tau)) = w_C(\tau)$. It immediately follows that $\tau = \tau_{opt}$ yields the

[14] In the case of certainty, the marginal likelihood of $\overline{\delta}$ occurring is infinitely large and hence, the probability of cooperation has to fall to zero as τ is increased beyond $\tau_{\max}$.

highest continuation value for $\tau \leq \widetilde{\tau}_{\max}$. Second, consider $\tau > \widetilde{\tau}_{\max}$. For $\tau \in (\widetilde{\tau}_{\max}, \widetilde{\tau}'_{\max}]$, τ is associated with $\eta(\tau) \in (\delta_{\min}, \delta_{\max})$. Suppose that $\eta(\tau_{opt}) \geq \delta_{\max}$ were a consistent solution for τ.(which it is not). Since $\eta(\tau) < \eta(\tau_{opt})$, it immediately follows from the proof of lemma 2 that $v_C(\tau, \eta(\tau)) < v_C(\tau, \eta(\tau_{opt}))$. Since $v_C(\tau, \eta(\tau_{opt})) = w_C(\tau) < w_C(\tau_{opt}) = v_C(\tau_{opt}, \eta(\tau_{opt}))$, it follows that $v_C(\tau, \eta(\tau)) < v_C(\tau_{opt}, \eta(\tau_{opt}))$. ■

In what follows the situation, in which the optimal degree of liberalization under certainty τ_{opt} cannot be attained with an associated likelihood of cooperation being chosen equal to unity (i.e. $\widetilde{\tau}_{\max} < \tau_{opt}$). In this situation the continuation value increases unambiguously in the degree of liberalization for $\tau \leq \widetilde{\tau}_{\max}$, because no deviations will ever occur and hence, there is no tradeoff between liberalization and sustainability of cooperation within this range of degrees of liberalization. But will the continuation value increase when the degree of liberalization is pushed beyond $\widetilde{\tau}_{\max}$ and the probability for breaching the agreement becomes strictly positive? The following proposition provides a sufficient condition for when increasing τ beyond $\widetilde{\tau}_{\max}$ is optimal.

Proposition 3 *When $\widetilde{\tau}_{\max} < \tau_{opt}$, there exists a threshold value $\overline{\varphi(\delta_{\max})}' \in (0, \overline{\varphi(\delta_{\max})}]$ such that a self-destructive agreement is preferred to a safe agreement ($\widetilde{\tau}_{opt} > \widetilde{\tau}_{\max}$) if*

$$\varphi(\delta_{\max}) < \frac{\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}{[\frac{1}{\delta_{\max}} + \frac{1}{\delta} - 1]\overline{\varphi(\delta_{\max})} + \frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}} \overline{\varphi(\delta_{\max})} \equiv \overline{\varphi(\delta_{\max})}'. \tag{7}$$

Proof. See the Appendix. ■

Hence, if the marginal likelihood of $\delta_{\max}$ occurring is sufficiently small, pushing liberalization beyond $\widetilde{\tau}_{\max}$ will be worthwhile.[15] To understand the intuition behind this result, it is important to emphasize the two opposing effects at work when τ is increased beyond $\widetilde{\tau}_{\max}$. Given that $\delta_{\max} > 1 - \tau_{opt}$ and hence $\widetilde{\tau}_{\max} < \tau_{opt}$, there is an unambiguously

[15] Note that it is implicitly assumed that it is not possible that v_C first decreases and then increases as τ increases beyond $\widetilde{\tau}_{\max}$. A sufficient condition for this not to be the case is that v_C is concave for $\tau \geq \widetilde{\tau}_{\max}$, something that can be shown to be true as long as w_D is not too convex in τ (i.e. if $\frac{d^2 w_D}{d\tau^2}$ is sufficiently small). If, for example, w_D is linear in τ, this will always be the case. Under such an assumption, an optimal solution exceeding $\widetilde{\tau}_{\max}$ will be possible if and only if v_C increases when τ marginally increases above $\widetilde{\tau}_{\max}$. If this assumption is relaxed, it is possible that $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, even if condition (7) is not satisfied.

positive effect on the outcome under mutual cooperation and thus, on the continuation value. But there is also a negative effect stemming from the increased ex ante likelihood of deviation occurring due to η falling below $\delta_{\max}$. This negative effect is directly related to the marginal likelihood of $\delta_{\max}$ occurring. The larger is this effect, i.e. the larger is $\varphi(\delta_{\max})$, the stronger will the negative impact of an increase in τ beyond $\widetilde{\tau}_{\max}$ be. Obviously, the negative effect is non-existent if $\varphi(\delta_{\max}) = 0$, and it will be worthwhile to increase τ beyond $\widetilde{\tau}_{\max}$ as long as there is a positive effect (which is the case when $\delta_{\max} > 1 - \tau_{opt}$).

The threshold value $\overline{\varphi(\delta_{\max})}'$ solely depends on $\overline{\delta}$ and $\delta_{\max}$; the term $\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}$ is, in fact, a function of $\delta_{\max}$. By plugging expression (6) into (7), an expression for $\overline{\varphi(\delta_{\max})}'$ can be obtained such that the impact of changes in $\overline{\delta}$ and $\delta_{\max}$ can be more easily assessed

$$\overline{\varphi(\delta_{\max})}' = \frac{1}{\frac{\frac{1}{\delta_{\max}} + \frac{1}{\overline{\delta}} - 1}{\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}} + (1 - \delta_{\max})(1 + \frac{\delta_{\max}}{\overline{\delta}})}. \tag{7'}$$

The most myopic realization $\delta_{\max}$ enters expression (7') both directly and indirectly through $\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}$. First, consider the shape of $\frac{\frac{d(w_C - w_N)}{d\tau}}{w_C - w_N}$. Since w_C is concave in τ, it immediately follows that $\frac{\frac{d(w_C - w_N)}{d\tau}}{w_C - w_N}$ is strictly decreasing in τ. Because $\frac{dw_C}{d\tau}|_{\tau=0} > 0$, $\frac{\frac{d(w_C - w_N)}{d\tau}}{w_C - w_N}$ goes to infinity as τ approaches zero, while $\frac{\frac{d(w_C - w_N)}{d\tau}}{w_C - w_N}|_{\tau=\tau_{opt}} = 0$. Since $\delta_{\max} > 1 - \tau_{opt}$, it follows that $\widetilde{\tau}_{\max} < \tau_{opt}$ and hence, $\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} > 0$. An increase in $\delta_{\max}$ leads to a lower $\widetilde{\tau}_{\max}$ and thus a higher value for $\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}$, its value equalling zero for $\delta_{\max} = 1 - \tau_{opt}$ and going to infinity as $\delta_{\max}$ approaches unity. Therefore, the first term of the denominator will unambiguously decrease when $\delta_{\max}$ is increased. The effect on the second term of the denominator is unambiguously positive if and only if $\delta_{\max} > \frac{1-\overline{\delta}}{2}$. Hence, if $\delta_{\max} \geq \frac{1-\overline{\delta}}{2}$, an increase in $\delta_{\max}$ unambiguously increases $\overline{\varphi(\delta_{\max})}'$.[16] Moreover $\overline{\varphi(1 - \tau_{opt})}' = 0$, while $\lim_{\delta_{\max} \to 1} \overline{\varphi(\delta_{\max})}' = \infty$. A low value of $\delta_{\max}$ implies a low $\overline{\varphi(\delta_{\max})}'$, and hence it will be worthwhile to increase τ beyond $\widetilde{\tau}_{\max}$ only for low marginal likelihoods of $\delta_{\max}$. A high value of $\delta_{\max}$ implies a high $\overline{\varphi(\delta_{\max})}'$

[16] Note that $\delta_{\max} \geq \frac{1-\overline{\delta}}{2}$ will always hold for $\overline{\delta} \geq \frac{1}{3}$. Note also that even if $\delta_{\max} < \frac{1-\overline{\delta}}{2}$, the effect of an increase in $\delta_{\max}$ on $\overline{\varphi(\delta_{\max})}'$ may be unambiguously positive (it will depend on the specification of w_C, however).

and hence, increasing τ beyond $\widetilde{\tau}_{\max}$ will be worthwhile except when the marginal likelihood of $\delta_{\max}$ occurring is very high; in fact it will always be beneficial to increase τ beyond $\widetilde{\tau}_{\max}$, if $\delta_{\max} = 1$, in which case $\widetilde{\tau}_{\max} = 0$ and no safe agreement with a strictly positive degree of liberalization is feasible.

The intuition is as follows. A higher $\delta_{\max}$ implies that the scope for safe agreements is smaller. The positive effect of increasing τ beyond $\widetilde{\tau}_{\max}$ thus increases in $\delta_{\max}$. The negative effect of an increase in τ beyond $\widetilde{\tau}_{\max}$ on the likelihood of cooperation, which depends on $\varphi(\delta_{\max})$, can thus be larger the larger is $\delta_{\max}$, without the overall impact of an increase in τ beyond $\widetilde{\tau}_{\max}$ becoming negative. For $\delta_{\max} = 1$, the positive effect of increasing τ above zero is infinitely large, thus always outweighing the negative impact on the likelihood of cooperation.

The effect of an increase in $\overline{\delta}$ on $\overline{\varphi(\delta_{\max})}'$ is unambiguously positive. The underlying reason is that a larger $\overline{\delta}$ implies a lower expected weight attributed to the future. Since the risk of breakdown occurring increases in the number of future periods, a lower expected weighting of the future implies a lower weighting of the negative impact of increasing τ beyond $\widetilde{\tau}_{\max}$. The following lemma summarizes the above findings.

Lemma 3 *The scope for self-destructive agreements being preferred to safe agreements, increases if*
(i) $\overline{\delta}$ is sufficiently large; and if
(ii) $\delta_{\max}$ is sufficiently large; and if
(iii) $\varphi(\delta_{\max})$ is sufficiently small.

It remains to establish how far τ should be pushed beyond $\widetilde{\tau}_{\max}$, when v_C is increasing as τ is increased above $\widetilde{\tau}_{\max}$. The following lemma provides an upper bound on the optimal choice of degree of liberalization.

Lemma 4 *If $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, then $\widetilde{\tau}_{opt} < \widetilde{\tau}'_{\max}$.*

Proof. See the Appendix. ■

Hence, when increasing τ beyond $\widetilde{\tau}_{\max}$ is optimal, pushing liberalization as far as possible will never be optimal.

6 CONCLUSIONS

The answer to the question posed in the title is affirmative: it may be optimal to agree on a degree of liberalization such that cooperation will eventually break down. A preference for self-destructive rather than safe agreements may arise in the present context of an infinitely repeated Prisoner's Dilemma tariff setting game between two symmetric countries, where the stochastically determined weight each government attributes to current vis-à-vis future payoffs is only privately observable.

The optimal agreement under certainty can be replicated under uncertainty if governments will always attribute sufficient weight to the future. Hence, applying the degree of liberalization that is optimal under certainty, while maintaining certainty of cooperation, will only be possible when the weight given to current payoffs is sufficiently low even under the most myopic realization of the random variable. If the latter is not the case, the degree of liberalization should at least be set at the level where the probability of deviation occurring just becomes strictly positive. Pushing the degree of liberalization further, the positive effect of more liberalization will have to be weighed against the negative impact of the ex ante likelihood of cooperation breaking down becoming strictly positive. If the latter outweighs the former, implementing the most far-reaching safe agreement is optimal. Else, it is optimal to implement a self-destructive agreement with a higher degree of liberalization than under the most far-reaching safe agreement. The latter outcome is more likely for a large ex ante expected weight given to current payoffs, for a large maximum possible weight attributed to current payoffs, and a small marginal likelihood of the maximum possible weight given to current payoffs.

There are, however, ways of overcoming the problem that agreements will not always be infinitely sustainable. Safeguard provisions similar to those found in trade agreements such as the GATT, for example, allow signatory countries to withdraw liberalization commitments under the agreement in order to protect certain overriding interests under specified conditions, thereby possibly eliminating the risk of breakdown.

7 References

Abreu, Dilip, Pearce, David, and Stacchetti, Ennio (1986), *Optimal Cartel Equilibria with Imperfect Monitoring*, Journal of Economic Theory, 39, 251-269

Athey, Susan and Bagwell, Kyle (2001), *Optimal Collusion with Private Information*, Rand Journal of Economics, 32(3), 428-465

Athey, Susan, Bagwell, Kyle, and Sanchirico, Chris (2004), *Collusion and Price Rigidity*, Review of Economic Studies, 71(2), 317-349

Bagwell, Kyle and Staiger, Robert W. (1990), *A Theory of Managed Trade*, American Economic Review, 80(4), 779-795

Baldwin, Richard (1987), *Politically Realistic Objective Functions and Trade Policy*, Economics Letters, 24, 287-290

Baye, Michael R. and Jansen, Dennis W. (1996), *Repeated Games with Stochstic Discounting*, Economica, 63, 531-541

Baye, Michael R. and Kovenock, Dan (2004), *Price Wars*, mimeo, Indiana University

Bernheim, B. Douglas and Dasgupta, Aniruddha (1995), *Repeated Games with Asymptotically Finite Horizons*, Journal of Economic Theory, 67, 129-152

Cramton, Peter C. and Palfrey, Thomas R. (1990), *Cartel Enforcement With Uncertainty About Costs*, International Economic Review, 31(2), 17-47

Dixit, Avinash (1987), *Strategic Aspects of Trade Policy*, in Bewley, T.F. (ed.), Advances in Economic Theory: Proceedings of the Fifth World Congress of the Economic Society, 329-362, Cambridge University Press, Cambridge, UK

Feenstra, Robert C. (1987), *Incentive Compatible Trade Policies*, Scandinavian Journal of Economics, 89(3), 373-387

Feenstra, Robert C. and Lewis, Tracy R. (1991), *Negotiated Trade Restrictions With Private Political Pressure*, Quarterly Journal of Economics, 106(4), 1287-1307

Green, Edward J. and Porter, Robert H. (1984), *Noncooperative Collusion under Imperfect Price Information*, Econometrica, 52(2), 87-100

Herzing, Mathias (2005), *Escape and Optimal Compensation in Trade Agreements*, mimeo, Institute for International Economic Studies, Stockholm University, Sweden

Hungerford, Thomas L. (1991), *GATT: A Cooperative Equilibrium in a Noncooperative Trading Regime*, Journal of International Economics, 31, 357-369

Jensen, Richard and Thursby, Marie (1990), *Tariffs with Private Information and Reputation*, Journal of International Economics, 29(1-2), 43-67

Kandori, Michihiro and Matsushima, Hitoshi (1998), *Private Observation, Communication and Collusion*, 66(3), 627-652

Kovenock, Dan and Thursby, Marie (1992), *GATT: Dispute Settlement and Cooperation*, Economics and Politics, 4(2), 151-170

Martin, Alberto and Vergote, Wouter (2005), *Antidumping: Cooperation Enhancing Retaliation?*, mimeo, Columbia University, New York

Mitchell, Matthew F. and Moro, Andrea (2004), *Informationally Efficient Trade Barriers*, mimeo, University of Iowa

Riezman, Raymond (1991), *Dynamic Tariffs with Asymmetric Information*, Journal of International Economics, 30(3-4), 267-284

Rotemberg, Julio J. and Saloner, Garth (1986), *A Supergame-theoretic Model of Price Wars During Booms*, American Economic Review, 76(3), 390-407

Spulber, Daniel F. (1995), *Bertrand Competition When Rivals' Costs Are Unknown*, Journal of Industrial Economics, 43(2), 1-11

Tower, Edward (1975), *The Optimum Quota and Retaliation*, Review of Economic Studies, 42, 623-630

Vives, Xavier (1984), *Duopoly Information Equilibrium: Cournot and Bertrand*, Journal of Economic Theory, 34, 71-94

8 Appendix

8.1 Derivation of v_D and v_C

$$v_D = \overline{\delta} w_N + (1-\overline{\delta})v_D \Leftrightarrow v_D = w_N \tag{3}$$

$$\begin{aligned}
v_C &= \Phi(\eta)\{\int_{\delta_{\min}}^{\eta} [\delta w_C + (1-\delta)v_C]d\Phi + \int_{\eta}^{\delta_{\max}} [\delta w_D + (1-\delta)v_D]d\Phi\} \\
&\quad + [1-\Phi(\eta)]\{\int_{\delta_{\min}}^{\eta} [\delta w_S + (1-\delta)v_D]d\Phi + \int_{\eta}^{\delta_{\max}} [\delta w_N + (1-\delta)v_D]d\Phi\} \\
&= \Phi(\eta)\{(w_C - w_D)\int_{\delta_{\min}}^{\eta} \delta d\Phi + w_D \int_{\delta_{\min}}^{\delta_{\max}} \delta d\Phi + (v_C - v_D)\int_{\delta_{\min}}^{\eta} (1-\delta)d\Phi\} \\
&\quad + [1-\Phi(\eta)]\{(w_S - w_N)\int_{\delta_{\min}}^{\eta} \delta d\Phi + w_N \int_{\delta_{\min}}^{\delta_{\max}} \delta d\Phi\} + v_D \int_{\delta_{\min}}^{\delta_{\max}} (1-\delta)d\Phi \\
&= \Phi(\eta)\{(w_C - w_D)\int_{\delta_{\min}}^{\eta} \delta d\Phi + \overline{\delta} w_D + (v_C - v_D)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi]\} \\
&\quad + [1-\Phi(\eta)]\{(w_C - w_D)\int_{\delta_{\min}}^{\eta} \delta d\Phi + \overline{\delta} w_N\} + (1-\overline{\delta})v_D \\
&\overset{(3)}{=} v_D + (w_C - w_D)\int_{\delta_{\min}}^{\eta} \delta d\Phi + (w_D - w_N)\overline{\delta}\Phi(\eta) \\
&\quad + (v_C - v_D)\Phi(\eta)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi] \\
\Leftrightarrow v_C &= w_N + \frac{\overline{\delta}\Phi(\eta) - \tau \int_{\delta_{\min}}^{\eta} \delta d\Phi}{1 - \Phi(\eta)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi]}(w_D - w_N).
\end{aligned} \tag{4}$$

8.2 Proof of Proposition 1

Given that there exist no more than three solutions to the CSE, and since $f(\delta_{\max}) = \delta_{\max}$ for $\tau = \widetilde{\tau}_{\max}$, it must be that $\lim_{\eta \to \delta_{\max}-} f'(\eta) < 1$ (note that $f'(\eta) = 0$ for $\eta > \delta_{\max}$) for $\tau = \widetilde{\tau}_{\max}$ for there to exist strictly positive solutions to the CSE when τ is increased beyond $\widetilde{\tau}_{\max}$. The first derivative of $f(\eta)$ for $\eta \in (\delta_{\min}, \delta_{\max})$ is calculated as follows

$$f'(\eta) = \frac{\begin{array}{c}[(\overline{\delta}-\tau)\Phi^2(\eta)+\tau][2\overline{\delta}\Phi(\eta)\varphi(\eta)-\tau\Phi(\eta)\varphi(\eta)\eta-\tau\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi] \\ -2(\overline{\delta}-\tau)\Phi(\eta)\varphi(\eta)[\overline{\delta}\Phi^2(\eta)-\tau\Phi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi]\end{array}}{\{\overline{\delta}\Phi^2(\eta)+\tau[1-\Phi^2(\eta)]\}^2}$$

$$= \frac{\begin{array}{c}(\overline{\delta}-\tau)\Phi^2(\eta)[-\tau\Phi(\eta)\varphi(\eta)\eta+\tau\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi] \\ +\tau[2\overline{\delta}\Phi(\eta)\varphi(\eta)-\tau\Phi(\eta)\varphi(\eta)\eta-\tau\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi]\end{array}}{\{\overline{\delta}\Phi^2(\eta)+\tau[1-\Phi^2(\eta)]\}^2}$$

$$= \tau\varphi(\eta)\frac{2\overline{\delta}\Phi(\eta)-\tau[\Phi(\eta)\eta+\int_{\delta_{\min}}^{\eta}\delta d\Phi]-(\overline{\delta}-\tau)\Phi^2(\eta)[\Phi(\eta)\eta-\int_{\delta_{\min}}^{\eta}\delta d\Phi]}{\{\overline{\delta}\Phi^2(\eta)+\tau[1-\Phi^2(\eta)]\}^2}.$$

Hence,

$$\lim{}_{\eta\to\delta_{\max}-} f'(\eta) = (1-\delta_{\max})(1+\frac{\delta_{\max}}{\overline{\delta}})\varphi(\delta_{\max}).$$

Thus, if $\varphi(\delta_{\max}) < \frac{1}{(1-\delta_{\max})(1+\frac{\delta_{\max}}{\overline{\delta}})} \equiv \overline{\varphi(\delta_{\max})}$, then $f'(\delta_{\max})|_{\tau=\widetilde{\tau}_{\max}} < 1$ and $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$.

8.3 Proof of Proposition 4

Since η unambiguously falls as τ is increased beyond $\widetilde{\tau}_{\max}$, it suffices to establish under what conditions $\frac{dv_C}{d\eta}|_{\tau=\widetilde{\tau}_{\max}} < 0$. The first derivative $\frac{dv_C}{d\eta}$ is given by

$$\frac{dv_C}{d\eta} = \frac{\partial v_C}{\partial \eta} + \frac{\partial v_C}{\partial \tau}\tau'(\eta).$$

In the proof of lemma 4, the CSE was plugged into the expression for v_C, given by (4), such that the following expression for v_C, multiplicatively separable in η and τ, was obtained. The partial derivatives of that expression, $\frac{\partial v_C}{\partial \eta}$ (see the proof of lemma 4) and $\frac{\partial v_C}{\partial \tau}$, are given by

$$\frac{\partial v_C}{\partial \eta} = \frac{\Phi^2(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi + \varphi(\eta)\eta^2}{\{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi\}^2}\overline{\delta}(w_D - w_N)$$
$$\frac{\partial v_C}{\partial \tau} = \frac{\overline{\delta}\eta\Phi(\eta)}{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi}\frac{d(w_D - w_N)}{d\tau}.$$

Thus,

$$\frac{\partial v_C}{\partial \eta}|_{\tau=\widetilde{\tau}_{\max}} = [1 + \frac{\delta_{\max}^2}{\overline{\delta}}\varphi(\delta_{\max})](w_D - w_N)|_{\tau=\widetilde{\tau}_{\max}}$$
$$\frac{\partial v_C}{\partial \tau}|_{\tau=\widetilde{\tau}_{\max}} = \delta_{\max}\frac{d(w_D - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}.$$

Rearranging the CSE yields an expression such that τ is a function of consistent solutions η

$$\eta = \Phi(\eta)\frac{\overline{\delta}\Phi(\eta) - \tau\int_{\delta_{\min}}^{\eta} \delta d\Phi}{\overline{\delta}\Phi^2(\eta) + \tau[1-\Phi^2(\eta)]}$$
$$\Leftrightarrow \tau = \frac{\overline{\delta}(1-\eta)\Phi^2(\eta)}{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi} \equiv \tau(\eta).$$

Hence,

$$\tau'(\eta) = \frac{\begin{array}{c}\overline{\delta}\{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi\}\{2(1-\eta)\Phi(\eta)\varphi(\eta) - \Phi^2(\eta)\} \\ -\overline{\delta}(1-\eta)\Phi^2(\eta)\{1-\Phi^2(\eta) - \Phi(\eta)\varphi(\eta)\eta + \varphi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi\}\end{array}}{\{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi\}^2}$$

$$= \overline{\delta}\Phi(\eta)\frac{\begin{array}{c}2(1-\eta)\eta[1-\Phi^2(\eta)]\varphi(\eta)+2(1-\eta)\Phi(\eta)\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi \\ -[1-\Phi^2(\eta)]\Phi(\eta)\eta-\Phi^2(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi-(1-\eta)\Phi(\eta)[1-\Phi^2(\eta)] \\ +(1-\eta)\Phi^2(\eta)\varphi(\eta)\eta-(1-\eta)\Phi(\eta)\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi\end{array}}{\{[1-\Phi^2(\eta)]\eta+\Phi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi\}^2}$$

$$= \overline{\delta}\Phi(\eta)\frac{\begin{array}{c}(1-\eta)\eta[2-\Phi^2(\eta)]\varphi(\eta)+(1-\eta)\Phi(\eta)\varphi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi \\ -\Phi(\eta)[1-\Phi^2(\eta)]-\Phi^2(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi\end{array}}{\{[1-\Phi^2(\eta)]\eta+\Phi(\eta)\int_{\delta_{\min}}^{\eta}\delta d\Phi\}^2}.$$

Thus,

$$\tau'(\eta)|_{\tau=\widetilde{\tau}_{\max}} = (1-\delta_{\max})(1+\frac{\delta_{\max}}{\overline{\delta}})\varphi(\delta_{\max})-1 = \frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}}-1.$$

Hence,

$$\frac{d(v_C-v_D)}{d\eta}|_{\tau=\widetilde{\tau}_{\max}} = [1+\frac{\delta_{\max}^2}{\overline{\delta}}\varphi(\delta_{\max})](w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}} + \delta_{\max}[\frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}}-1]\frac{d(w_D-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}.$$

Thus,

$$\frac{d(v_C-v_D)}{d\eta}|_{\tau=\widetilde{\tau}_{\max}} < 0 \Leftrightarrow \frac{\frac{d(w_D-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}}} > \frac{1+\frac{\delta_{\max}^2}{\overline{\delta}}\varphi(\delta_{\max})}{\delta_{\max}[1-\frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}}]}.$$

Since $w_D - w_N = \frac{w_C-w_N}{1-\tau}$, the term $\frac{\frac{d(w_D-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}}}$ can be rewritten in the following way:

$$\begin{aligned}\frac{\frac{d(w_D-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}}} &= \frac{\frac{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}}{(1-\widetilde{\tau}_{\max})^2}}{(w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}}} + \frac{\frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{1-\widetilde{\tau}_{\max}}}{(w_D-w_N)|_{\tau=\widetilde{\tau}_{\max}}} \\ &= \frac{1}{1-\widetilde{\tau}_{\max}} + \frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}} \\ &= \frac{1}{\delta_{\max}} + \frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}}.\end{aligned}$$

Hence,

$$
\begin{aligned}
&\frac{d(v_C - v_D)}{d\eta}|_{\tau=\widetilde{\tau}_{\max}} < 0 \\
&\Leftrightarrow \frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} > \frac{1 + \frac{\delta_{\max}^2}{\overline{\delta}}\varphi(\delta_{\max})}{\delta_{\max}[1 - \frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}}]} - \frac{1}{\delta_{\max}} \\
&\Leftrightarrow \delta_{\max}[1 - \frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}}]\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} > \frac{\delta_{\max}^2}{\overline{\delta}}\varphi(\delta_{\max}) + \frac{\varphi(\delta_{\max})}{\overline{\overline{\varphi(\delta_{\max})}}} \\
&\Leftrightarrow \begin{bmatrix} \frac{\delta_{\max}^2}{\overline{\delta}} + \frac{1}{\overline{\overline{\varphi(\delta_{\max})}}} \\ + \frac{\delta_{\max}}{\overline{\overline{\varphi(\delta_{\max})}}}\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} \end{bmatrix} \varphi(\delta_{\max}) < \delta_{\max}\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} \\
&\Leftrightarrow \varphi(\delta_{\max}) < \frac{\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}{\frac{\delta_{\max}}{\overline{\delta}}\overline{\varphi(\delta_{\max})} + \frac{1}{\delta_{\max}} + \frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}\overline{\varphi(\delta_{\max})} \\
&\qquad = \frac{\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}{[\frac{\delta_{\max}}{\overline{\delta}} + \frac{(1-\delta_{\max})(1+\frac{\delta_{\max}}{\overline{\delta}})}{\delta_{\max}}]\overline{\varphi(\delta_{\max})} + \frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}\overline{\varphi(\delta_{\max})} \\
&\qquad = \frac{\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}{[\frac{1}{\delta_{\max}} + \frac{1}{\overline{\delta}} - 1]\overline{\varphi(\delta_{\max})} + \frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}}}\overline{\varphi(\delta_{\max})} \equiv \overline{\varphi(\delta_{\max})}'.
\end{aligned}
$$

Given that $\delta_{\max} > 1 - \tau_{opt}$, it follows that $\widetilde{\tau}_{\max} < \tau_{opt}$ and hence, $\frac{\frac{d(w_C - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C - w_N)|_{\tau=\widetilde{\tau}_{\max}}} > 0$, whereby it immediately follows that $\overline{\varphi(\delta_{\max})}' \in [0, \overline{\varphi(\delta_{\max})}]$.

8.4 Proof of Lemma 6

From the proof of proposition 4, we know that

$$\frac{dv_C}{d\eta}|_{\tau=\widetilde{\tau}'_{\max}} = \frac{\partial v_C}{\partial \eta}|_{\tau=\widetilde{\tau}'_{\max}} + \frac{\partial v_C}{\partial \tau}|_{\tau=\widetilde{\tau}'_{\max}}\tau'(\eta)|_{\tau=\widetilde{\tau}'_{\max}}.$$

It is easily established that $\frac{\partial v_C}{\partial \eta} > 0$ and $\frac{\partial v_C}{\partial \tau} > 0$ for any $\eta > \delta_{\min}$. Since $\widetilde{\tau}'_{\max}$ is implicitly defined by $\tau'(\eta) = 0$, it immediately follows that $\frac{dv_C}{d\eta}|_{\tau=\widetilde{\tau}'_{\max}} > 0$, which implies that $\frac{dv_C}{d\tau}|_{\tau=\widetilde{\tau}'_{\max}} < 0$. Hence, $\widetilde{\tau}_{opt} < \widetilde{\tau}'_{\max}$.