The GATT and Gradualism[1]

Ben Lockwood, *University of Warwick*

Ben Zissimos, *Vanderbilt University*[2]

First version: April 2002

Revised February 2004

ABSTRACT: This paper shows how the institutional rules imposed on its signatories by the GATT created a strategic incentive for countries to liberalize gradually. Free trade can never be achieved if punishment for deviation from a trade agreement is limited to a 'withdrawal of equivalent concessions.' Trade liberalization must be gradual if, in addition, deviation from an agreement is limited. The paper shows how (sufficiently patient) countries may have an incentive to deviate in a limited way when operating under GATT dispute settlement procedures.

KEYWORDS. Nash tariffs, Free Trade, Gradualism, Trade agreement.
JEL CLASSIFICATION NUMBERS: F02, F13, F15, C73.

[1]An earlier version of this paper was circulated under the title 'Trade Agreements with Limited Punishments.' We would like to thank two anonymous referees, Rick Bond, Andrew Daughety, Jayasri Dutta, Jennifer Reinganum, Robert Staiger, Quan Wen, John Whalley and Kar-yiu Wong for helpful comments and conversations about this paper. We are also grateful for comments from seminar participants at the Midwest International Meeting, University of Wisconsin, and The WTO and World Trade II Meeting, University of Washington. Support from the ESRC's Evolving Macroeconomy Programme, The Warwick Centre for Public Economics, and The Centre for the Study of Globalization and Regionalization, University of Warwick is gratefully acknowledged.

[2]Corresponding author: Dept of Economics, Vanderbilt University, Nashville, TN 37235.
Tel: ++1 615 322 3339.
E-mail: ben.zissimos@vanderbilt.edu

1. Introduction

The experience of trade liberalization in the period since World War II has presented economists with two puzzles. First, even in developed countries, free trade has remained stubbornly elusive, with average trade-weighted tariffs remaining at low but still positive levels.[3] Second, tariffs have been cut only gradually under the General Agreement of Tariffs and Trade (GATT). Since the GATT was drawn up after the war, tariffs have fallen from a trade weighted average of 50 percent to around 5 percent today. Neither of these two facts sits well with the textbook view that sees a trade agreement as a simple repeated Prisoner's Dilemma: that is, as a situation where it is individually rational for countries to impose tariffs, but collectively rational to abolish them.

The purpose of this present paper is to propose an explanation for these two puzzles by focusing on two particular aspects of the rules imposed on trade liberalization by the GATT. We examine the implications for the trade liberalization process of the dispute settlement system under the GATT, particularly a *withdrawal of equivalent concessions* (WEC). In doing so, we show that the incentives created by these rules were sufficient to motivate the outcomes of failure to reach free trade and gradual liberalization actually observed.

Our work responds to concerns raised by Bagwell and Staiger (1999) about the effectiveness of the two 'GATT pillars,' reciprocity and nondiscrimination, to guide governments from inefficient unilateral outcomes to the efficiency frontier. Indeed, the basic approach that we take to modelling these issues was established by Bagwell and Staiger, who were the first to characterize and analyze the fundamental features of the GATT in a game theoretic, general equilibrium framework.[4] Bagwell and Staiger's (1999) central point builds on the observation that the main economic purpose of a trade agreement

[3]It could be argued that there is no puzzle in the failure to reach free trade. In a world where trade carries externalities, current positive tariff levels could be efficient. However, in practice there appears to be a consensus that efficiency has not been reached; that mutual gains from trade are still available from further multilateral trade liberalization. To keep things simple, free trade will be used as a metaphor for this 'yet to be obtained' efficient level of international trade.

[4]Elements of the analytical framework for analysis of the GATT that Bagwell and Staiger (1999) construct can be traced to their earlier papers published in 1990, 1996 and 1997. For a synthesis, which introduces, develops and extends their approach to a comprehensive treatment of the GATT/WTO as an economic institution, see Bagwell and Staiger (2002).

is to overcome the terms of trade externality that occurs when countries can influence the world price through policy interventions. They argue that the two GATT pillars of reciprocity and nondiscrimination provide a mechanism by which the terms of trade externality can be overcome. But Bagwell and Staiger also point out that in practice enforcement difficulties at the international level may preclude governments from fully eliminating the terms-of-trade driven restrictions in trade volumes and arriving at the frontier. Our two main results, that free trade cannot be reached, and that the level of openness that is possible can be approached only gradually, provide substantive analytical support for Bagwell and Staiger's concerns.[5]

The focus of this paper is on the broad sweep of trade liberalization under the GATT in the post war period, up to the conclusion of the Uruguay Round in 1994. The idea is to assume that two countries have signed the GATT, and then analyze the dynamic equilibrium (liberalization) path that results when a tariff reduction game is played according to GATT rules. We undertake a formal representation of the GATT Articles in question, setting them out in a fully specified (game theoretic, general equilibrium) framework. This includes a formal statement of the GATT rules that govern deviation from an agreement and the degree of retaliation allowed. We also formalize the GATT's stipulations as to the treatment of countries that break the GATT rules themselves. Thus, not only are we able to characterize the liberalization process when liberalization takes place according to the Articles. We are also able to show that, having signed up to the GATT, countries could do no better than liberalize according to the Articles.[6] This property of the theoretical framework accords well with the historical period from the GATT's inception in 1947 to

[5] Developing their earlier (1999) discussion, Bagwell and Staiger (2002 Chapter 6) take up the issue of enforcement difficulties. Their general discussion is wide ranging and comprehensive, touching on some of the issues that we consider here. But formally they consider a different set of issues. They look at a situation where efficiency cannot be enforced because governments are not sufficiently patient, and then show how 'escape clauses' under Article XIX of the GATT can rebalance an agreement, relaxing the incentive to deviate. This they class as 'on-equilibrium-path' retaliation.

By contrast, we study 'off-equilibrium-path' retaliation, where Articles XIX and XXIII sanction varying severities of retaliation, depending on the degree of deviation from the equilibrium path. Bagwell and Staiger point out that aspects of Articles XIX and XXIII sanction both on-equilibrium-path and off-equilibrium-path retaliation, depending on the circumstance. Thus the discussion we undertake in the present paper and that of Bagwell and Staiger (2002 Chapter 6) are complementary in that they consider different ways in which Articles XIX and XXIII can make a trade agreement self-enforcing. See Section 4 for further details and discussion.

[6] In formal terms, we show in all cases that the efficient equilibrium path is subgame perfect, and that given a deviation the punishment path is also subgame perfect.

the conclusion of the Uruguay Round. Over that time, the GATT Articles were adhered to quite closely. For example, violations of tariff bindings were not often observed; see Chapter 2 of Whalley and Hamilton (1996) for further details.[7]

The characterization of WEC is as follows. Suppose that a deviant country fails to implement some agreed market access measure. Under GATT rules, contracting parties to the agreement are allowed to do no more than to withdraw market access concessions equivalent to those that the deviant failed to implement. We model exactly this penalty structure in the context of a dynamic game and examine its implications for trade liberalization under the GATT.

Our first main result is that the WEC rule does facilitate trade liberalization but, when retaliation is limited by the WEC rule, free trade certainly cannot be reached no matter how little countries discount the future. This result contrasts markedly with conventional insights from the theory of repeated games, which indicate that free trade can be achieved, given sufficiently little discounting. The intuition behind this first main result is simple. A standard repeated game allows trade partners to implement the worst (credible) punishment against a deviant. In general, the WEC rule makes such severe punishments illegal. By outlawing a class of severe punishments, the WEC rule compromises efficiency. But there is an additional subtlety to the result that is worth noting. The generality of the result follows from the fact that under WEC, countries choose the severity of their own punishment by the extent of their initial deviation. Therefore, for any discount rate there always exists a worthwhile deviation from free trade, which is smaller for lower discount rates, given that future punishments will be no worse than the initial deviation. Note that for this first result, partial irreversibility is imposed only on one side of the agreement. That is to say, WEC limits only the actions of punishers and not deviators from an agreement.

When only the actions of a punisher are restricted we find that gradualism does not arise; only when the actions of an initial deviant are restricted as well does gradualism

[7] In 1994, as part of the conclusion to the Uruguay Round, signatories to the GATT formed the World Trade Organization (WTO). To some extent the analysis of the present chapter is relevant for the period since 1994 too, because the GATT Articles were adopted in the Charter of the WTO (GATT 1994). But since the WTO's inception, we appear to be observing a change in the operation of the regime, with a number of instances where rules have been broken. The reasons for this change present an important agenda for future research, but will not be taken up here in this present paper.

occur. We argue that a restriction on initial deviations arises from the variation in possible punishment that results. Following Bagwell and Staiger (1990, 2002 Chapter 6), in response to a relatively small deviation WEC is allowed. However, under Article XXIII contracting parties may also authorize, in 'appropriately serious cases,'[8] a 'suspension of GATT obligations.' This, Bagwell and Staiger argue, may be associated with a standard trigger-strategy punishment by reversion to myopic best response tariffs.

Our second main result incorporates these restrictions on deviants and, in so doing, characterizes gradualism. We show that if punishments are constrained by the WEC rule *and* the initial deviation by any country is also constrained, then the most efficient self enforcing path of trade liberalization is gradual. Because punishment is limited, current tariff cuts can only be made self-enforcing by the promise of future tariff reductions. But if initial deviation is sufficiently limited as well, then it is always possible to promise liberalization over a number of future periods that would more than compensate. So on the equilibrium path, trade liberalization must take place over a number of periods.

We have already mentioned that Bagwell and Staiger's work has established the general framework in which we operate. Setting this framework in a wider context, the paper builds on a substantial literature going back to Johnson (1953-54).[9] Early contributions explain trade liberalization in a standard repeated game framework, where tariff cuts from their one-shot Nash equilibrium values are explained as the outcome of self-enforcing trigger strategies (Dixit 1987).[10] As remarked above, using a trigger strategy has two limitations. It cannot explain gradualism and moreover, free trade is always a self-enforcing outcome with sufficiently little discounting.

More recent literature has offered several explanations as to why self-enforcing tariff

[8] There is no formula for the determination of 'appropriately serious cases.' However, for the purposes of formalization we make precise the circumstances in which a suspension of GATT obligations occurs in the model. Details and discussion are provided in the text, especially in Section 4.

[9] Horwell (1966), and more recently Lockwood and Wong (2000) compare trade wars with specific and ad valorem tariffs, showing the outcomes to be different under the respective instruments. Hamilton and Whalley (1983) broaden considerably the basis on which tariff wars can be examined by showing how they can be studied using numerical simulations.

[10] Among many others, some contributions to the literature on trade agreements that use the threat of retaliation as threat points in cooperative or non-cooperative models include Mayer (1981), Bagwell and Staiger (1990), Bond and Syropoulos (1996) and McLaren (1997). Syropoulos (2002) examines the effect of country size, showing that if one trade partner is larger than another by a significantly large ratio, then it will prefer a trade war to a free trade agreement.

agreements are gradual. The general idea is that, initially, full liberalization cannot be self-enforcing, because the benefits of deviating from free trade are too great to be dominated by any credible punishment. But if there is partial liberalization, structural economic change within the domestic economy reduces the benefits of deviation from further trade liberalization (and/or raises the costs of punishment to the deviator). The individual papers differ in their description of the structural change induced by partial liberalization. Staiger (1995) endows workers in the import competing sector with specific skills, making them more productive there than elsewhere in the economy. When they move out of this sector they lose their skills with some probability, relaxing the constraint on further trade liberalization. In Devereux (1997), there is dynamic learning-by-doing in the export sector. In Furusawa and Lai (1999), there are linear[11] adjustment costs incurred when labor moves between sectors. Bond and Park (2002) show that gradualism arises as a result of asymmetry in country size. In Chisik (2003), increasing sunk costs of investment in the expanding export sector raise the costs of deviation, and increased specialization gradually lowers the lowest obtainable self-enforcing tariff. All of these papers focus on elements of the domestic economy to motivate gradualism, as opposed to elements of the international trading system that we study here.

This present paper also makes a wider contribution to the applied game theory literature on gradualism. In particular, Lockwood and Thomas (2002) study the effect of *complete* irreversibility of strategic actions in an abstract game, showing that irreversibility on the side both of the initial deviator and the punisher are sufficient for gradualism. In this present paper we extend the framework of Lockwood and Thomas from an abstract game to a tariff game. Having obtained new results in the tariff game framework, we can extend the insights of the present paper to the wider field of applied game theory by thinking of tariffs in a tariff game as an example of strategic actions in an abstract game. Then in the first part of this present paper we would say that *partial* irreversibility of the strategic action (the tariff) is assumed on the side of the punisher, but the initial deviation is unrestricted. We then see explicitly that gradualism cannot result. Only when there is a degree of irreversibility on both sides does gradualism arise. In this sense,

[11] Furusawa and Lai have an Appendix where they show that with strictly convex adjustment costs, a social planner would choose gradual tariff reduction.

the present paper extends Lockwood and Thomas (2002).[12]

The paper proceeds as follows. The next section sets up the basic analytical framework, defining formally the tariff reduction game and the punishments for deviation allowed under the GATT, focusing primarily on WEC. Section 3 then studies trade agreements under WEC. It is here that we will see how trade liberalization is achieved in this framework but that free trade cannot be reached. Section 4 then examines trade agreements when both the extent of initial deviations and punishments is limited. It is under these circumstances that gradual trade liberalization takes place. Section 5 concludes.

2. Optimal Tariffs, Trade Agreements and Punishments under GATT

2.1. Tariffs and Welfare

We work with a standard model of international trade in which two countries produce and consume two final goods.[13] Country i is assumed to have a comparative advantage in the production of commodity i. Countries are symmetrical in all other respects. In particular, preferences and technologies are identical subject to a re-ordering of the goods. Both countries are large enough to affect the terms of trade. Import tariffs are the only form of trade restriction allowed; τ_t^i represents the tariff set by i on imports from j in period t. Preferences of the representative consumer in each country over both goods are given by a strictly quasiconcave utility function. The two countries' production possibility loci are strictly concave to the origin over the two goods.[14] Without loss of generality, a

[12] In a two player prisoners' dilemma with continuous actions, under *complete irreversibility* once players have achieved a given level of cooperation neither can reverse their action in order to punish the other. Under *partial irreversibility*, some reversal of actions is possible. Lockwood and Thomas study only complete irreversibility on the side of both deviator and punisher while we study partial irreversibility on the side of just the punisher as well as on the side of both deviator and punisher.

[13] As Bagwell and Staiger (1999) point out, a two country model is sufficient to carry out an examination of the limits to reciprocal trade liberalization which is the subject here. A model incorporating more countries is only required when issues of non-discrimination against third parties are under discussion. We do analyze an n country model in Lockwood and Zissimos (2001), and show that the main results hold, but we do not establish subgame perfection of punishments under GATT with respect to trigger strategies in that earlier paper. Establishing subgame perfection for the case of n countries would be difficult and it is not clear that insight would be added.

[14] This general specification also encompasses 'endowment models of international trade.' All the general analysis of the paper is worked out for an endowment model example in the appendix.

country does not set tariffs on its own goods. Also note that $-1 < \tau_t^i < \infty$.

Within a period, $t = 1, 2, \ldots$, the order of events is as follows. First, each country i, observing the tariffs set by the other country (and their own) up to the previous period, simultaneously chooses an import tariff. Then, given world prices of the two goods and tariffs, perfect competition in production takes place. Next, the representative consumer in each country chooses consumption to maximize utility subject to budget constraints. This yields the usual indirect utility function and excess demands. Then, conditional on tariffs, markets clear and world prices for the goods are determined.[15] This world price will of course depend on tariffs, as will tariff revenues. We assume that equilibrium prices are unique.

So, we can write equilibrium welfare of any country i as a function of tariffs only. Let τ be the tariff that a country levies on its imports, and let τ' be the tariff that its exports face upon entry to the foreign market. In equilibrium, the indirect utility function can be written $w(\tau, \tau')$. Now we can define a *Nash equilibrium in tariffs* in the usual way as a $\hat{\tau}$ such that $w(\hat{\tau}, \hat{\tau}) \geq w(\tau, \hat{\tau})$ all $\tau \in (-1, \infty)$, $i \in \{1, 2\}$. We will focus on symmetric Nash equilibria. Such equilibria exist and are unique for the special cases that we consider below, due to the symmetry of the model.

As we are focussing on tariff reductions, we will assume throughout that $(\tau, \tau') \in [0, \widehat{\tau}]^2 = F^2$. We assume three properties of w:

A1. $w_1(\tau, \tau') \geq 0$, $w_2(\tau, \tau') \leq 0$, for all $(\tau, \tau') \in F^2$, and $w_1(\tau, \tau') > 0$ if $\tau < \hat{\tau}$, $w_2(\tau, \tau') < 0$ if $0 < \tau'$.

A1 asserts that whenever the other country's tariff is below Nash equilibrium, a country likes an increase in its own tariff, and a reduction in the tariff of the other country. In other words, the static tariff game has a Prisoner's Dilemma structure. Our second assumption is very weak:

A2. $w_1(\tau, \tau) + w_2(\tau, \tau) < 0$ for all $(\tau, \tau) \in F^2$ with $\tau > 0$.

This says that any equal reduction in all tariffs, starting from a situation of equal tariffs

[15] As this is a general equilibrium model, prices are determined only up to a scalar, and so some normalization (e.g. choice of numeraire) must be made.

at or below the Nash level, makes any country better off. Moreover, note that from the optimality of free trade, $w_1(0,0)+w_2(0,0) = 0$. Our third assumption is:

A3. $w_{12}(\tau, \tau') < 0$, all $(\tau, \tau') \in F^2$.

That is, tariffs are strategic substitutes; the closer the other country's tariff is to the Nash equilibrium tariff, the smaller the gain a country makes from increasing its own tariff.

2.2. Trade Agreements

We are interested in how fast countries can reduce tariffs from this non-cooperative Nash equilibrium, and also whether they can ever reach free trade i.e. $\tau_t^1 = \tau_t^2 = 0$, if the tariff reduction plan must be *self-enforcing* i.e. the outcome of a subgame-perfect equilibrium. Payoffs over the infinite horizon are discounted by a common discount factor δ, $0 < \delta < 1$ i.e.

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t w(\tau_t^i, \tau_t^j). \tag{2.1}$$

A *tariff history* at time t is defined as a complete description of all past tariffs in both countries; $h_t = \{(\tau_1^1, ..., \tau_{t-1}^1, \tau_1^2, ..., \tau_{t-1}^2)\}$. Both countries can observe tariff histories. A *tariff strategy* for country $i = 1, 2$ is defined as a choice of tariffs τ_t^i in periods $t = 1, 2...$ conditional on every possible tariff history. A *tariff path* of the game is a sequence $\{(\tau_t^1, \tau_t^2)\}_{t=1}^{\infty}$ that is generated by the tariff strategies of both countries.

Given the symmetry of the model, we restrict our attention to *symmetric* equilibrium[16] tariff paths where $\tau_t^i = \tau_t$, $t = 1, 2, \ldots$, i.e. where both countries choose the same tariff in every time period, and we denote such paths by the sequence $\{\tau_t\}_{t=1}^{\infty}$.

2.3. Punishments under the GATT

Suppose that $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ is a candidate for an equilibrium tariff sequence, where $\widetilde{\tau}_t$ is the tariff "agreed" for period t. In a standard repeated tariff game, the punishment that i levies on j for deviating from $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ is to raise its tariff to the Nash level $\hat{\tau}$ and maintain it there, the most severe credible punishment (which we call a *trigger strategy*).

[16] In the sequel, it is understood that "equilibrium" refers to subgame-perfect Nash equilibrium.

In practice, GATT signatories were bound by Article XVIII to adopt a *withdrawal of equivalent concessions* (WEC). Under the WEC rule country i, upon observing that j has deviated at time $t-1$, withdraws precisely the equivalent concessions to market access at time t. That is, if the deviant j has set $\tau^j_{t-1} = \tau' > \widetilde{\tau}_{t-1}$, where $\tau' < \hat{\tau}$, then in the next period instead of retaliating by setting $\hat{\tau}$ the other party withdraws the concessions made, implementing $\tau^i_t = \tau' = \tau^j_{t-1}$ as well. We wish to analyze equilibrium paths that result when countries are bound by the WEC rule as opposed to standard trigger strategies.

We will show that once a country has deviated by setting $\tau' > \widetilde{\tau}_{t-1}$, and the other country has invoked punishment τ' under the WEC rule, then maintaining tariffs at τ' thereafter is a subgame perfect equilibrium. In principle, one country could deviate subsequently by setting $\tau'' > \tau'$, in which case the other country will again punish by WEC in the period after deviation by setting τ''. But we show that under the present assumptions no country has an incentive to act in this way (Proposition 1).

In addition, we must specify what would happen if a country did not adhere to WEC when it punished another country for deviation. Under Article XXIII contracting parties are given the power, 'in appropriately serious cases, to authorize a contracting party or parties to suspend GATT obligations to other contracting parties' (Jackson 1989, page 94). In the present stylized framework, we say that to *break WEC*, by setting $\tau^i_t > \tau' = \tau^j_{t-1} > \widetilde{\tau}_{t-1}$, results in an *indefinite suspension of GATT obligations* among both parties; there is an indefinite return to trigger strategies $\hat{\tau}$.[17] We show that WEC is a subgame perfect punishment strategy given the alternative of breaking WEC and provoking an indefinite suspension of GATT obligations (Proposition 2).[18]

[17] By breaking WEC, Bagwell and Staiger (2002 Chapter 6) argue that a government 'ploughs over the final backstop of a GATT ruling' which will lead to a break-down of cooperation, formally equated to trigger strategies.

[18] In the paper we follow the relatively simplistic approach of assuming that punishments are imposed symmetrically. This assumption has been criticized because all parties have an incentive to 'renegotiate away from' these punishments during the punishment phase. Farrell and Maskin (1989) and van Damme (1989) show that punishments must be asymmetric in order for the punisher to do better under punishment than any feasible alternative. However, following Blume (1994) and McCutcheon (1997) we argue that symmetric punishments can be used to enforce a trade agreement if renegotiation away from punishments sanctioned by the GATT is sufficiently costly. Such renegotiation of the rules is outside the authority of trade negotiators and is more costly than negotiation 'within the rules,' requiring ratification in national legislatures rather than routine rubber stamping by trade negotiators. McCutcheon shows that when renegotiation is more costly than the stipulated trigger strategies then trigger strategies may be used to sustain an agreement. Renegotiation proofness requires finite punishments. Our use of infinite punishments is a simplifying but inessential short-cut.

3. Failure to Reach Free Trade

In this section we analyze the trade agreement that is feasible when only the actions of punishers are restricted by WEC. The actions of initial deviants are completely unrestricted, as in a standard repeated tariff game. We will see that on the efficient equilibrium path some liberalization is possible but that free trade cannot result. This result is shown to hold in general for all discount rates. However, the efficient liberalization path does not exhibit gradualism. For that, the actions of initial deviants must be restricted as well (as in Section 4).

3.1. Optimal Deviations

We begin by characterizing the optimal deviation from a symmetric equilibrium path $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ for any country i, given that it rationally anticipates that it will be punished by the WEC rule. Let i's optimal deviation at t from the reference path $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ be denoted z_t. Note that the withdrawal of equivalent concessions applies only to deviation by setting a tariff *above* the agreed rate $\widetilde{\tau}_t$. Under WEC, no punishment is imposed if a country deviates by cutting tariffs *below* the agreed level $\widetilde{\tau}_t$. Thus there is an asymmetry in the penalty. Formally, the payoff any country can expect from a deviation to z_t is:[19]

$$\Delta(z_t, \{\widetilde{\tau}_t\}_{t=1}^{\infty}) = \begin{cases} (1-\delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t) & \text{if } z_t > \widetilde{\tau}_t \\ (1-\delta)w(z_t, \tilde{\tau}_t) + (1-\delta)\sum_{s=t+1}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t) & \text{if } z_t < \widetilde{\tau}_t \end{cases} \tag{3.1}$$

We are interested in the *optimal deviation* z_t i.e. the choice of z_t that maximizes $\Delta(z_t, \{\widetilde{\tau}_t\}_{t=1}^{\infty})$ given the reference path. The largest possible gain from deviation is the supremum of $\Delta(z_t, \{\widetilde{\tau}_t\}_{t=1}^{\infty})$ across all values of $z_t \neq \widetilde{\tau}_t$, which we denote by $\overline{\Delta}(\{\widetilde{\tau}_t\}_{t=1}^{\infty})$.

Lemma 1. *Assume A1-A2. Then,*

$$\overline{\Delta}(\{\widetilde{\tau}_t\}_{t=1}^{\infty}) = \max\{\max_{z_t \geq \tilde{\tau}_t} [(1-\delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t)],\ (1-\delta)\sum_{s=t}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t)\}.$$

This result says that the best that a country can do is either to replicate the payoff on the equilibrium path - the second term in curly brackets - or to deviate by setting tariffs

[19] We show in Proposition 1 below that (3.1) generates a punishment path that is subgame perfect.

above the agreed level; $z_t \geq \tilde{\tau}_t$. It can never benefit by a unilateral deviation $z_t < \tilde{\tau}_t$.[20] Now, from the first term in curly brackets which gives the gains to deviation, define

$$z(\tau_t) = \arg\max_{z_t \geq \tilde{\tau}_t} \{(1-\delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t)\}. \tag{3.2}$$

$z(.)$ can be thought of as a kind of "reaction function" indicating how the optimal deviation varies with the agreed tariff $\tilde{\tau}_t$. We can now obtain a characterization of $z(.)$ that is very useful. Define

$$\zeta(\tau) = \arg\max_{z} \{(1-\delta)w(z, \tau) + \delta w(z, z)\} \tag{3.3}$$

This is the solution to the problem in equation (3.2), ignoring the inequality constraint. Note that

$$\zeta'(\tau) = \frac{(1-\delta)w_{12}(z,\tau)}{D}$$

where $D > 0$ from the second-order condition for the choice of z in (3.3). So if A3 holds then $\zeta'(\tau) < 0$. Also, define $\overline{\tau}$ to satisfy:

$$\overline{\tau} = \zeta(\overline{\tau}) \tag{3.4}$$

This is a self-enforcing tariff level: i.e. at $\overline{\tau}$ the optimal deviation is in fact not to deviate at all.

We now have the following characterization of $z(.)$:

Lemma 2. *Assume A1-A3. There is a unique solution to (3.4), for which $\overline{\tau} < \widehat{\tau}$. The solution to (3.2) satisfies: (i) for all $\tau < \overline{\tau}$, $z(\tau) = \zeta(\tau) \geq \overline{\tau} > \tau$; (ii) for all $\tau \geq \overline{\tau}$, $z(\tau) = \tau$; (iii) $z(0) > 0$.*

Lemma 2 shows that for all tariffs τ less than $\overline{\tau}$ the optimal deviation is to set a tariff above $\overline{\tau}$; $z(\tau) = \zeta(\tau) \geq \overline{\tau} > \tau$. For tariffs τ greater than or equal to $\overline{\tau}$ the optimal deviation is just τ itself; $z(\tau) = \tau$. Recall that $\zeta(\tau)$ is downward sloping and goes through the point $\overline{\tau}$. Therefore, $\zeta(\tau) < \tau$ for all $\tau > \overline{\tau}$. But then the constraint $z_t \geq \tau_t$ in (3.2) is binding.

[20] To see why, recall that a withdrawal of equivalent concessions applies only to upward deviations. If a country were to deviate by setting a tariff that were lower than agreed - $z_t < \tau_t$ - the WEC rule would not require all other countries to follow the deviant downwards. We can therefore ignore the possibility that $z_t < \tau_t$ because, by A1, a country would make itself worse off by deviating in this way.

We now have a complete characterization of the optimal deviation z_t, given any tariff τ_t. So for any $\widetilde{\tau}_t$ in a candidate equilibrium sequence $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ we know the optimal deviation for that period under WEC. This will be used to characterize uniquely the efficient equilibrium path.

Before moving on to look at equilibrium paths, it remains to confirm that (3.1) generates a punishment path that is subgame perfect, and that a withdrawal of equivalent concessions is itself a subgame perfect punishment strategy. These two desired properties are established in the following two results.

Proposition 1. *Assume A1-A3. The payoff to deviation given by* (3.1) *generates a punishment path that is subgame perfect.*

Proposition 1 confirms that once a country has deviated to $z(\tau)$ and the other country has retaliated by also adopting $z(\tau)$, neither country has an incentive to deviate again under WEC. To see why Proposition 1 holds, remember that WEC does not require downwards deviations to be matched.[21] This feature of the rule is captured formally by the constraint $z_t \geq \tau_t$ in (3.2). At first sight it might appear that an initial deviation to $z(\tau_t)$ could create the incentive for a further deviation to $z(z(\tau_t)) > z(\tau_t)$, once the punisher had also adopted $z(\tau_t)$. In fact, under our assumptions about the model's structure, this would never happen, and there is no incentive to deviate subsequently from $z(\tau_t)$. To see why, first note that in general an initial deviation may occur at $z(\tau_t) > \overline{\tau}$, or $z(\tau_t) = \overline{\tau}$. (Lemma 2 rules out the possibility that $z(\tau_t) < \overline{\tau}$.) If $z(\tau_t) > \overline{\tau}$ then, from (3.3), we have that $\zeta(\tau_t) < \overline{\tau}$; without the constraint $z_t \geq \tau_t$ the best response would be to lower and not to raise tariffs. But the constraint $z_t \geq \tau_t$ does bind in (3.2), and so the best response to $z(\tau_t)$ is itself. And if $z(\tau_t) = \overline{\tau}$ then the best response once again is $z(\tau_t)$ because $\overline{\tau}$ is self-enforcing. Once a country has deviated to $z(\tau)$, we must confirm that the other country actually has an incentive to adopt a punishment consistent with WEC, and that after that neither country has a further incentive to deviate. Confirmation is obtained in the next result.

Proposition 2. *Assume A1, A2. Then WEC is a subgame perfect punishment strategy.*

[21] Recall that the rule applies only to a *withdrawal* and not an *extension* of equivalent concessions.

The alternative would be to punish more severely than is allowed by WEC and trigger a suspension of GATT obligations under Article XXIII. In the light of Proposition 1, the result of Proposition 2 is not surprising. First note that whether WEC is adhered to or broken, $z(\tau_t)$ is the initial deviation in both cases. But the optimal deviation $z(\tau_t)$ is chosen to maximize the present discounted payoff, rationally anticipating that $z(\tau_t)$ will also be the subsequent punishment. Under a suspension of GATT obligations, while the initial deviation is also $z(\tau_t)$, the subsequent punishment is fixed (at $\hat{\tau}$). So a deviation that breaks WEC, τ', does not take future retaliation into account.[22] Therefore τ' must be suboptimal relative to $z(\tau_t)$, given that calculation of $z(\tau_t)$ does take into account the subsequent continuation payoffs.

Given conventional repeated game logic, this result might at first appear surprising. We are used to the idea that if a country is very impatient, at the limit $\delta = 0$, then it will always gain by reneging on an agreement, triggering $\hat{\tau}$. But under WEC a country's impatience is reflected in its choice of $z(\tau_t)$; to pick the severest case, if $\delta = 0$ it is easy to see that $z(\tau_t) = \hat{\tau}$ because the problem collapses to a stage game. So there is nothing to gain under deviation from WEC (because under deviation the payoff would also be determined by $\hat{\tau}$), even under extreme impatience. It is then not surprising that the result holds for all $\delta \in (0, 1)$.

3.2. Efficient Equilibrium Paths and Failure to Reach Free Trade

We can now formally define the conditions that must hold if a symmetric tariff path is to be a subgame-perfect one in our game. In every period, the continuation payoff from the path must be at least as great as the maximal payoff from deviation, given that a punishment consistent with the WEC will ensue. From Lemma 1, the maximal relevant payoff from deviation at t is $(1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t))$. So, formally, we require:

$$(1-\delta)(w(\tau_t, \tau_t) + \delta w(\tau_{t+1}, \tau_{t+1}) + ... \quad) \geq (1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t)), \; t = 1, ... \tag{3.5}$$

Of course, a whole set of paths will satisfy this sequence of inequalities: let this set of equilibrium paths be denoted E. An *efficient tariff reduction path* in the set E is simply a

[22] More specifically, the first order condition in the derivation of the deviation tariff τ' does not take future retaliation into account because $\hat{\tau}$ is a constant in the objective function.

sequence $\{\tau_t\}_{t=1}^{\infty}$ of tariffs in E for which there is no other sequence $\{\tau'_t\}_{t=1}^{\infty}$ also in E which gives a higher payoff to any country, as calculated by (2.1). Following the arguments of Lockwood and Thomas (2002), it can be shown that if $\{\tau_t\}_{t=1}^{\infty}$ is efficient, (3.5) holds with equality at every date i.e. :

$$(1-\delta)(w(\tau_t,\tau_t)+\delta w(\tau_{t+1},\tau_{t+1})+... \quad) = (1-\delta)w(z(\tau_t),\tau_t)+\delta w(z(\tau_t),(z(\tau_t)),\ t=1,... \tag{3.6}$$

The intuition is that if (3.5) held with strict inequality, it would be possible to reduce the tariff path by a small amount without violating (3.5).

Of the class of equilibrium paths E, it is obviously the efficient path (shown to be unique below) that is of most interest. It is generally accepted that the GATT provides a mechanism through which countries are able to coordinate their selection of the efficient path (Bagwell and Staiger 1990). We now turn to characterizations of the efficient equilibrium path. Our first main result, Proposition 3, establishes that free trade is in fact impossible under WEC.

Proposition 3. *(Failure to reach free trade.) Assume A1-A3. Let $\{\tau_t\}_{t=1}^{\infty}$ be an equilibrium path. Then $\tau_t > 0$, for all $\delta < 1$, all t.*

The proof of this Proposition works by showing that if one country adopts free trade at *any* point in time, then the other will have an incentive to deviate by levying a positive tariff. So such an agreement would not be self-enforcing. The result follows from Lemma 2, which shows that the best response to free trade is a positive tariff. This is clearly in contrast to the standard case with unlimited punishments. For in that case, countries can credibly punish deviators by reverting to (for example) Nash tariffs, and then it is well-known that for some $\delta_0 < 1$, free trade can be attained in equilibrium for all $\delta > \delta_0$. In the present case, by contrast, the extent of the punishment is endogenously determined by WEC to be exactly as severe as the initial deviation. So there always exists a worthwhile deviation, decreasing with the discount rate, given that under WEC the punishment can do no more than match the initial deviation in the subsequent periods.

We now turn to the more difficult question of what form the efficient path takes. Say that an equilibrium tariff reduction path is a *stationary* path if $\tau_t = \tau$, all $t \geq 1$ (recall $\tau_0 = \hat{\tau}$); that is, there is an immediate and permanent tariff reduction. A stationary equilibrium path must satisfy:

$$\alpha(\tau) \equiv \max_{z \geq \tau}\{(1-\delta)w(z,\tau) + \delta w(z,z)\} \leq w(\tau,\tau) \equiv \beta(\tau).$$

To characterize such paths, note first the properties of α, β. First, β is decreasing in τ by A2, and α is decreasing by A1, A2. Second, at the Nash equilibrium, as $z = \hat{\tau}$ is a best response to $\hat{\tau}$, $\alpha(\hat{\tau}) = \beta(\hat{\tau})$ i.e. the non-cooperative Nash equilibrium is a stationary equilibrium path[23]. Third,

$$\alpha(0) \equiv \max_{z \geq 0}\{(1-\delta)w(z,0) + \delta w(z,z)\} > w(0,0) \equiv \beta(0)$$

as a small increase in z from 0 strictly increases $w(z,0)$ (from A1), while leaving $w(z,z)$ unchanged (as $w(z,z)$ is maximized at zero, by A2).

So, the possibilities are shown in Figure 1. Next, as α, β are both downward-sloping, they may have multiple crossing-points, as shown. Note that $\alpha(\tau)$ and $\beta(\tau)$ coincide over the range $\overline{\tau} \leq \tau \leq \hat{\tau}$. This is because, by Lemma 2, $z(\tau) = \tau$ for all $\tau \geq \overline{\tau}$. So

$$\begin{aligned} \alpha(\tau) &= \max_{z \geq \tau}\{(1-\delta)w(z,\tau) + \delta w(z,z)\} \\ &= w(\tau,\tau) = \beta(\tau) \text{ for all } \tau \geq \overline{\tau} \end{aligned}$$

Finally, the *smallest* stationary equilibrium tariff will be at the lowest crossing point of α, β, namely τ^*. Moreover, using Lemma 2, it is possible to show that under some additional assumptions, $\tau^* = \overline{\tau}$. Formally, we have:

Proposition 4. *Assume A1-A3. Let $\tau_0 = \hat{\tau}$. There is a unique efficient stationary path, $\tau_t = \tau^*$, all $t \geq 1$, where $\tau^* > 0$ is the smallest root of the equation $\alpha(\tau) = \beta(\tau)$. Moreover, if A3 holds, and $w_{11}(\tau,\tau)$, $w_{22}(\tau,\tau) \leq 0$ on $[0,\tau]$, then $\tau^* = \overline{\tau} < \hat{\tau}$.*

Proposition 4 shows that under a withdrawal of equivalent concessions it is possible for both countries to agree to reduce tariffs immediately to the level $\overline{\tau}$, holding them there indefinitely, and moreover, this is the best equilibrium stationary path. The result is illustrated in Figure 2, which refines Figure 1.

[23] Note that it is *not* claimed that $\hat{\tau} = \zeta(\hat{\tau})$. In fact, it is easily checked from the definition of (3.3) that $\zeta(\hat{\tau}) < \hat{\tau}$, so the constraint $z \geq \hat{\tau}$ in the definition of α binds, implying that $z(\hat{\tau}) = \hat{\tau}$, and consequently, that $\alpha(\hat{\tau}) = (1-\delta)w(\hat{\tau},\hat{\tau}) + \delta w(\hat{\tau},\hat{\tau}) = w(\hat{\tau},\hat{\tau}) = \beta(\hat{\tau})$.

The question then arises as to whether there is a *non-stationary* path in E which is more efficient than the stationary path $\tau_t = \overline{\tau}$, $t \geq 1$. The following result answers this negatively:

Proposition 5. *Assume A1-A3. The stationary path, which has* $\widetilde{\tau}_t = \overline{\tau}$*, all* $t \geq 1$*, is the unique efficient path in E.*

The idea of the proof is the following. If there is a more efficient equilibrium path, then it must involve a tariff $\tau_t < \overline{\tau}$. But, the dynamics of (3.6), expressed as a difference equation, tell us that once $\tau_t < \overline{\tau}$, $\tau_{t+1} < \tau_t$ i.e. the path must be monotonically decreasing. But this is impossible, as either it implies a stationary equilibrium path below $\overline{\tau}$ (impossible by definition), or a tariff sequence diverging to minus infinity (which cannot be efficient).

Our results can be illustrated with a quasi-linear example. Preferences can be assumed to take the form of (A.14). This example is analyzed thoroughly in the appendix. In the appendix we first show that the Nash tariff is $\hat{\tau} = 1/(\sigma - 1)$. It is then shown that

$$\overline{\tau} = \frac{1-\delta}{\sigma(1+\delta) - 1}. \tag{3.7}$$

Note from (3.7) that in general, $0 < \overline{\tau} < \hat{\tau}$. That is, $\overline{\tau} \to \hat{\tau}$ as $\delta \to 0$, and $\overline{\tau} \to 0$ as $\delta \to 1$. When agents place a high weight on future outcomes, tariff rates close to zero can be achieved under WEC. The elasticity of substitution between goods is also inversely related to the level of $\overline{\tau}$. In the appendix we also verify that (A.14) satisfies assumptions A1-A3.

If the GATT provides a means by which countries select the efficient tariff reduction path, then Propositions 3, 4 and 5 provide a complete characterization of this path. Accordingly, under WEC trade liberalization can be achieved, but free trade cannot be reached. However, at present our model cannot "explain" the gradualism in tariff-cutting observed in the post war period.

4. Gradualism

In this section we analyze the effect on equilibrium trade liberalization of imposing a degree of tariff irreversibility not just on the punisher but on the initial deviant as well. Recall that in the previous section partial irreversibility was imposed only on the punisher; the initial deviant was completely unrestricted in tariff setting.

Our discussion draws on Chapter 6 of Bagwell and Staiger (2002) where enforcement of international trade agreements under GATT is discussed.[24] As explained above, Bagwell and Staiger distinguish two types of initial deviation from an agreed tariff level; on-equilibrium-path deviations and off-equilibrium-path deviations. According to their description an off-equilibrium-path deviation, familiar from the theory of repeated games, is a reversion to the myopic best response tariff level. An on-equilibrium-path deviation is a smaller deviation required to keep the incentive compatibility constraint of the agreement binding in response to an unexpected surge in import volumes. Retaliation to both types of deviation is allowed for under the dispute settlement provisions of the GATT, in particular Article XXIII.

In response to a relatively small on-equilibrium-path deviation, a 'rebalancing of concessions' is allowed. This is where a trade partner simply withdraws concessions that it made to the deviator that were not reciprocated. Such a WEC can take place either under Article XIX if the measures can be agreed upon between parties themselves, or under Article XXIII if a panel is required to provide independent arbitration to help resolve a dispute in the interpretation of GATT rules.

However, under Article XXIII contracting parties may also authorize, in 'appropriately serious cases,' a 'suspension of GATT obligations.' This, Bagwell and Staiger argue, may be associated with a standard trigger-strategy punishment by reversion to myopic best response tariffs. In their discussion, they suggest that a suspension of GATT obligations would occur only once a panel ruling had been violated, either by a deviator that raised tariffs after a panel had ruled against them, or by a country that retaliated even when the initial complaint was not upheld. However, in their formalization, a relatively large (off-equilibrium-path) deviation is met by an immediate return to trigger strategies

[24] A more rigorous treatment of the model in Chapter 6 is presented in Bagwell and Staiger (1990).

while a relatively small (on-equilibrium-path) deviation is met by immediate WEC.

Although we present an alternative formalization, our model shares the same feature as Bagwell and Staiger's (2002 Chapter 6) that a relatively large deviation is met by an immediate return to trigger strategies while a relatively small deviation is met by immediate WEC. The difference is that in our formalization a small deviation and retaliation by WEC is an off-equilibrium-path deviation as well i.e. once a small deviation and WEC takes place there is no return to the equilibrium path.

The fact that WEC is an off-equilibrium-path deviation in our framework represents a difference in emphasis rather than a contradiction to Bagwell and Staiger's approach. Indeed, Bagwell and Staiger acknowledge that there is an off-equilibrium-path element to such a deviation because the deviator stands in violation of the agreement if it does not bring its policy back into conformity with GATT rules once the panel ruling is issued (see Bagwell and Staiger 2002, footnote 5 on page 98). Bagwell and Staiger emphasize the written form of the GATT, particularly of Article XIX, to argue that such deviations are temporary and WEC 'legalizes' the deviation by establishing a rebalancing of concessions. We, however, place emphasis on the observation by Dam (1970 page 100) that "most of the tariff increases made under Article XIX have in fact never been rescinded." Dam then goes on to point out that "an affected trade partner could always demand that the concession be reinstated and may invoke the dispute settlement procedures if no action is taken." The fact that in the majority of cases no such action was taken suggests that relatively small deviations were simply matched with WEC, exactly as in our formalization.[25]

There is a second sense in which our formalization of a smaller deviation and retaliation by WEC differs from Bagwell and Staiger's. While Bagwell and Staiger back such deviation by a stochastic shock, our modelling environment is stationary. Thus, our deviation and subsequent WEC is admittedly more difficult to justify. But again, Bagwell and Staiger identify two components to such a deviation. While their formalization emphasizes the response to the stochastic shock, they also point out that an exception to an agreement under Article XIX 'raises the prospect that a government may be motivated in part by a desire to shift the costs of its intervention onto its trading partner, thus

[25] Of course, according to the theory such off-equilibrium-path deviations should never have actually been observed, but could be justified as 'trembling' or 'learning' about the way the GATT rules worked.

upsetting ... the balance of concessions. In its original formulation, GATT's Article XIX addresses this possibility by allowing that the trading partner can then take a retaliatory exception and withdraw its own substantially equivalent concession.' (See Bagwell and Staiger 2002 page 105.) Thus Bagwell and Staiger clearly suggest that while part of a deviation is motivated directly by a shock, part is motivated by the desire to (opportunistically) shift costs. But presumably, providing the violation is not too big, the exception will be taken 'in good faith' and will simply be responded to by WEC. While Bagwell and Staiger formalize the part of a deviation that is in direct response to a shock, we formalize the part that is opportunistically motivated by a terms-of-trade gain.[26]

A natural question raised by Bagwell and Staiger's (2002) discussion and our formalization is, in the absence of (full) verification of a shock, how much a government can deviate and still have its action taken 'in good faith.' There is no basis written into GATT Articles on which to discriminate between a deviation that brings about a withdrawal of equivalent concessions and one that provokes a suspension of GATT obligations. But for the purposes of our analysis we need to formalize the distinction. To do this, say that along the candidate equilibrium path, at period $t-1$, a given tariff level τ_{t-1} is achieved. The aim is then to achieve some lower tariff level $\widetilde{\tau}_t$. At period $t-1$, the target tariff level $\widetilde{\tau}_t$ for period t is then said to be the *binding* for period t or *scheduled binding* (Article II).[27] By contrast, bindings already achieved for $t-s$, $s \geq 1$ are said to be *past bindings.*

Country i *breaks a scheduled binding* by setting $\tau_t = \tau'$ such that $\tau' > \widetilde{\tau}_t$; in period t it fails to reduce its tariff to $\widetilde{\tau}_t$. The punishment imposed on country i by country j then depends on the extent to which the binding is broken. We distinguish two possibilities.

(i) In period t, country i breaks a binding but does not raise its tariff above the level set in $t-1$; $\tau_{t-1} \geq \tau_t = \tau' > \widetilde{\tau}_t$. It does not break a past binding. Then in period $t+1$ country j withdraws equivalent concessions.

[26] The 'opportunistically motivated' part does not arise in Bagwell and Staiger's analysis because they analyze a symmetric situation so the terms of trade effects cancel; both countries deviate symmetrically because they both face the same shock, which both can verify. We capture an asymmetric situation in which only one country deviates, and its motives cannot be perfectly verified.

[27] In practice, a tariff binding has come to be understood to represent one of three types of commitment: (1) to lower a tariff (duty) to a stated level; (2) not to raise a tariff above its current level; (3) not to raise a tariff above a specified higher level (Dam 1970, page 31). In this present paper we focus on the most popular usage of the term binding given by (1).

(ii) In period t, country i *breaks a past binding* and does raise its tariff above the level set in $t-1$; $\tau_t > \tau_{t-1}$. Then in all future periods there is an indefinite suspension of GATT obligations; country j sets $\hat{\tau}$ from $t+1$ onwards.[28]

We say that when the actions of initial deviants are restricted in this way then we have a *tariff game with bindings.*

One aspect of this formalization may appear to go against the written procedure for dispute settlement and needs clarification. As mentioned previously the procedures of Article XXIII center on the notion of "nullification and impairment" and do not require the actual breach of a legal obligation (i.e. breaking a scheduled or past binding.) However, as Dam (1970 page 360) points out, dispute settlement panels came to favor "an approach that would make the legality of the trade measure under the substantive provisions of the General Agreement the crucial factor in determinations of nullification and impairment." According to this reading of events, through practice panels came to regard the breach of a legal obligation to be the key condition for nullification and impairment, as in our formalization (Jackson 1989 backs this view; see page 95).

To summarize, our behavioral assumption essentially implies that, even if it is not backed by a stochastic shock, an 'opportunistically motivated' deviation is taken in good faith and is simply matched by a WEC as long as it does not break a past binding.

We now show that if countries are sufficiently patient then they will not break a past binding under deviation in the knowledge that doing so will bring about a suspension of GATT obligations.

[28] Thus we introduce a behavioral assumption about what affected parties will tolerate and not an assumption based on GATT rules. But given that such a distinction is called for by Bagwell and Staiger, drawing the line at a past binding seems as reasonable as any. And it is entirely consistent with the situation formalized by Bagwell and Staiger where a large but verifiable shock causes a past binding to be broken but is nevertheless redressed by WEC.

It may nevertheless be argued that in practice parties have tolerated a breach of past bindings, and responded by WEC. Our model could be extended to allow a past binding to be broken by a certain margin $\lambda\tau_{t-1}$, where $\lambda > 1$, provided that $\lambda\tau_{t-1} << \hat{\tau}$. However, the analysis on which this extension is based is more complex, without changing the gradualism result qualitatively. Given, by Lemma 2, that all $z(\tau) > \overline{\tau}$ for $\tau < \overline{\tau}$, for the extension we would need an analogous $\vec{\tau}$ (say) for which $z(\vec{\tau}) = \lambda\vec{\tau}$. By construction it must be the case that $\vec{\tau} < \overline{\tau}$. Then $z(\tau) > \vec{\tau}$ for $\tau < \vec{\tau}$. Because deviation from $\vec{\tau}$ must be more attractive than deviation from $\overline{\tau}$, there must be an upper bound on the value of λ and the range of feasible $\delta < 1$ must be reduced.

Proposition 6. *Assume A1-A3 and that governments play a tariff game with bindings. Then there is a* $0 \leq \underline{\delta} < 1$ *such that for all* $\delta \in (\underline{\delta}, 1)$ *if a country deviates from a candidate equilibrium path* $\{\widetilde{\tau}_t\}_{t=1}^{\infty}$ *it does not break a past binding.*

This result is based on Friedman's (1971) "Nash-threats" folk theorem. A country knows that if it deviates from a candidate equilibrium path in such a way that it breaks a past binding, by setting $\tau_t = \tau' > \tau_{t-1}$, then this will provoke a suspension of GATT obligations. As with the "Nash-threats" folk theorem, this triggers an indefinite imposition of myopic best responses $\hat{\tau}$. On the other hand if a country deviates in such a way that it does not break a past binding, by setting $\tau' \leq \tau_{t-1}$, then under the punishment phase tariffs are τ'. And by A2, symmetric tariffs at τ' yield a higher payoff than symmetric tariffs at $\hat{\tau}$. At the limit $\delta = 1$, countries are hurt more by the lower payoff under suspension of GATT obligations $\hat{\tau}$. So there must exist a range $\delta \in (\underline{\delta}, 1)$ for which countries do not break past bindings under deviation.

How much does the bound on δ limit the applicability of this result and the others that rely on it? It is generally recognized that as long as agents are able to adjust the interval between periods, they can ensure that δ is in the required range for the result to hold when it is in their interest. (See Scherer 1980 and Fudenberg and Tirole 1991 for further discussion.) So a consistent explanation of the fact that past bindings were rarely broken (until 1994) is that dispute resolution procedures operated sufficiently quickly.

We can now reformulate the equilibrium conditions (3.5) under a tariff game with bindings. It is clear that in the event that a country deviates, the "optimal" deviation given in (3.2) will be chosen, unless $z(\tau_t) > \tau_{t-1}$, in which case τ_{t-1} will be chosen. So, defining $\chi(\tau_t, \tau_{t-1}) = \min\{z(\tau_t), \tau_{t-1}\}$, the equilibrium conditions become

$$\begin{aligned} (1-\delta)(w(\tau_t,\tau_t) + \delta w(\tau_{t+1},\tau_{t+1}) + \ldots \quad) &\geq \\ (1-\delta)w(\chi(\tau_t,\tau_{t-1}),\tau_t) + \delta w(\chi(\tau_t,\tau_{t-1}),\chi(\tau_t,\tau_{t-1})),\ t &= 1,\ldots \end{aligned} \tag{4.1}$$

As before, let the set of equilibrium tariff paths be E, and define the efficient tariff paths in E as those paths that maximize (2.1). Also as before, any efficient path must satisfy (4.1) with equality.

Proposition 6 says that there exists a range of δ for which a country's deviation will not exceed τ_{t-1}. It will be helpful in what follows to know the conditions under which the optimal deviation will be precisely equal to τ_{t-1}. We now establish in the next result that for $\tau_t, \tau_{t-1} < \overline{\tau}$, the optimal deviation is indeed equal to τ_{t-1}.

Lemma 3. *Assume A1-A3, let $\delta \in (\underline{\delta}, 1)$, and let $\overline{\tau}$ be the (unique) value for which $z(\tau) = \zeta(\tau) = \tau$ given δ. If $\tau_t, \tau_{t-1} \leq \overline{\tau}$, then $\chi(\tau_t, \tau_{t-1}) = \tau_{t-1}$.*

Recall that by Lemma 2, for all $\tau < \overline{\tau}$, $z(\tau) = \zeta(\tau) \geq \overline{\tau} > \tau$. If $\tau_t, \tau_{t-1} \leq \overline{\tau}$ then it follows immediately that $\chi(\tau_t, \tau_{t-1}) = \tau_{t-1}$. Then for all t, τ_{t-1} can be substituted for $\chi(\tau_t, \tau_{t-1})$ in the equilibrium condition (4.1), making characterization of an equilibrium path possible.

We take two steps to prove that the efficient tariff reduction path is unique and gradually decreasing under the tariff game with bindings. We start by assuming, in Lemma 4, that $\tau_t, \tau_{t-1} \leq \overline{\tau}$ for all $t \geq 1$ and show that the equilibrium tariff sequence must be strictly decreasing. Then in the final result of the paper, Proposition 7, we show that in the efficient equilibrium tariff sequence it must be the case that $\tau_1 < \overline{\tau}$. Gradual tariff reductions then follow by Lemma 4. We first present the results then give the intuition.

From (4.1) and Lemma 3, an efficient path with $\tau_t \leq \overline{\tau}$, $t \geq 1$ must satisfy[29]

$$w(\tau_t, \tau_{t+1}) = \frac{1}{\delta}\left[w(\tau_{t-1}, \tau_t) - w(\tau_t, \tau_t)\right] + \frac{w(\tau_{t-1}, \tau_{t-1})}{1-\delta} - \frac{\delta w(\tau_t, \tau_t)}{1-\delta}, \ t > 1. \tag{4.2}$$

Let $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ be the sequence that solves (4.2) with initial conditions τ_0, τ_1. We can now establish gradualism by showing that as long as there is a tariff reduction in the first period then tariffs must strictly fall in all subsequent periods along any efficient equilibrium path.

Lemma 4. *Assume A1 and $\delta \in (\underline{\delta}, 1)$. Any sequence $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ that satisfies (4.2), with initial conditions τ_0, τ_1 that satisfy $0 < \tau_1 < \tau_0$ and $0 < \tau_1 < \overline{\tau}$, is strictly decreasing i.e. $0 < \tau_{t+1}(\tau_0, \tau_1) < \tau_t(\tau_0, \tau_1)$ all $t \geq 2$.*

Now consider the construction of an efficient path, given these results. First, τ_0 is given at $\hat{\tau}$. Second, from $t = 2$ onwards, i.e. conditional on τ_0, τ_1, the unique efficient path is simply $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ as long as (i) $\tau_1 < \tau_0$ (required by Lemma 4), and (ii) $\tau_1 \leq \overline{\tau}$

[29] See the proof of Lemma 4 for a full derivation of (4.2).

(required by Lemma 3; otherwise, the efficient path does not satisfy (4.2)). So, it remains to choose $\tau_1 \leq \overline{\tau} < \hat{\tau}$. If the path is to be efficient, the incentive constraint (4.1) must hold with equality in period 1 i.e.

$$\begin{aligned} & (1-\delta)(w(\tau_1,\tau_1) + \delta w(\tau_2(\hat{\tau},\tau_1),\tau_2(\hat{\tau},\tau_1)) + ...) \\ = \; & (1-\delta)w(\chi(\tau_1,\hat{\tau}),\tau_1) + \delta w(\chi(\tau_1,\hat{\tau}),\chi(\tau_1,\hat{\tau})) \end{aligned} \tag{4.3}$$

We now have:

Proposition 7. *Assume A1-A3, $\delta \in (\underline{\delta}, 1)$, and that governments play a tariff game with bindings. There exists a smallest value of τ_1, $0 < \tilde{\tau}_1 < \overline{\tau}$, that satisfies (4.3). Consequently, the path $(\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3, ...)$ is the unique efficient path, with $\tilde{\tau}_t = \tau_t(\hat{\tau}, \tilde{\tau}_1)$, $t > 1$. This path exhibits a gradually decreasing tariff i.e. $\tilde{\tau}_{t+1} < \tilde{\tau}_t$, $t \geq 1$.*

From Proposition 7 we learn that under a tariff game with bindings it *is* possible to achieve an equilibrium path for which $\tilde{\tau}_t < \overline{\tau}$, all $t \geq 1$. Consider some period s in which tariffs have been reduced by a gradual process over periods $t = 1, ..., s$ to some tariff level $\tilde{\tau}_s < \overline{\tau}$. Now suppose that the agreement requires $\tilde{\tau}_{s+1} < \tilde{\tau}_s$ in period $s+1$. If the agreement proposes no further reductions in periods after that, then country j does better by deviating whilst country i proceeds to set $\tilde{\tau}_{s+1} < \tilde{\tau}_s$, even if country i imposes the WEC penal code in all periods after that. But by Proposition 6, country j does not break a past binding under deviation; its optimal deviation is to maintain $\tilde{\tau}_s$ in $s+1$.[30] And by Proposition 7, it is always possible for country i to promise additional reductions in future periods that can more than compensate for the gains to deviation in period $s+1$. This is gradualism in other words.

5. Conclusions

This paper helps to explain two stylized facts about trade liberalization, namely failure to reach free trade and gradualism, by studying the interplay between countries' unilateral incentive to set tariffs and aspects of the institutional structure set up in the framework of the GATT to achieve trade liberalization. In particular, aspects of the GATT's Articles

[30] This is conditional upon $\delta \in (\underline{\delta}, 1)$.

are shown to impose partial irreversibility on countries' ability to set tariffs. First, a withdrawal of equivalent concessions limits the ability of contracting parties to raise tariffs in order to punish a country for deviating from an agreement under Article XVIII. Second, while relatively small deviations may be punished by a withdrawal of equivalent concession, a larger deviation will provoke a suspension of GATT obligations. (As explained above, this second feature is conditional upon an behavioral assumption about parties' tolerance of deviations and how they are assumed to use available GATT rules as a result.) Partial irreversibility imposed by a withdrawal of equivalent concessions is sufficient to rule out free trade on the efficient equilibrium path, but not to motivate gradualism. In order for gradualism to arise on the efficient equilibrium path, initial deviations must be partially irreversible as well.

The starting point of our analysis is to take the GATT rules that we study as given and assume that two countries have each already signed the GATT. We then play out a dynamic tariff reduction game and characterize the equilibrium path. Off-equilibrium-path play is fully characterized by drawing on aspects of Articles XIX and XXIII. In so doing we are able to show that (sufficiently patient) countries can do no better than to keep to a symmetrical gradual tariff reduction path.

We show that once countries have signed up to the GATT, they can do no better than to liberalize according the rules that it sets out. Yet it is clear that the GATT would achieve greater efficiency if it sanctioned more severe punishments of deviators. This raises an important question for future research, namely, what features are omitted from the model in this present paper that might make WEC and the possibility to deviate under Articles XIX and XXIII achieve efficiency? One direction in which to pursue an answer would be to introduce stochastic productivity shocks to our model. Because it is deterministic, our model rules out the possibility that Article XIX will actually be used, as it was originally intended, to relax the participation constraint in a trade agreement in the face of shocks that would otherwise precipitate the agreement's break-up and a return to greater protectionism. Whilst our analysis shows that WEC in conjunction with Article XIX prevents efficiency from being achieved, in a stochastic world WEC might be efficiency enhancing by enabling an agreement to survive. The efficiency enhancing potential of Articles XIX and XXIII is discussed by Bagwell and Staiger (1990, and 2002

Chapter 6). The interplay of these competing roles for deviation and retaliation presents an interesting agenda for future research.

Inevitably, our theoretical framework simplifies the situation in a number of other key respects. Countries are assumed to be symmetrical, each country exports only a single good, with both countries equally open at a given time. In practice countries export a number of goods, with levels of openness varying across sectors. Variation in country size and purchasing power across different markets is likely to make the actual dynamics of liberalization considerably more subtle and complex. Gradualism in a context where there are asymmetries across countries has been studied by Bond and Park (2000), but not within the context of the GATT penalty structure that we examine here.

A promising direction for future research would allow trade block formation to be considered. The theory of repeated games has been used to study trade block formation, where a preferential trade agreement is supported by the credible threat of punishment. In a recent paper using a repeated game framework Bond, Syropoulos and Winters (2001) point out that trade liberalization within the European Union has been very slow.[31] It may be that our framework could be adapted to provide a way of understanding gradualism between members.

There may be many other competing pressures other than the standard terms-of-trade motive working against further liberalization, and these are also suppressed in our model. One area that has attracted significant attention recently is the incentive for politicians to give in to protectionist inducements from interest groups (Grossman and Helpman 1995). These protectionist forces may have been outweighed at an early stage in the post-war trade liberalization process when liberalization gains were large relative to the rents from protectionism, but not later once the potential trade gains began to be exhausted. Future research could usefully study the interaction of these counteracting forces.

The main point of the present paper is that under GATT rules trade liberalization must be gradual. A natural question follows as to 'how gradual' trade liberalization becomes as a result. Given the competing explanations for gradualism, it would be difficult

[31] Bond, Riezman and Syropoulos (2003) also draw attention to the gradual nature of liberalization within FTAs.

to disentangle which is attributable to the forces for gradualism that we describe here. One possibility is to use simulations. In an earlier version of this present paper, Lockwood and Zissimos (2002), we made some progress in this direction by undertaking simulations based on the quasi-linear preferences presented in the appendix. Yet these were unsatisfactory in that they appeared to show 'too much liberalization too soon' given what we have actually observed. Further work is needed to establish how the forces for gradualism vary with different functional forms and whether they tend to exacerbate other factors that slow down the liberalization process.

A. Appendix

A.1. Proof of Propositions

Proof of Lemma 1. (a) First, suppose that a country deviates to $z_t < \widetilde{\tau}_t$. Then, from (3.1), as there is no retaliation, future payoffs are unaffected by the choice of deviation. Moreover, as $w(z_t, \tilde{\tau}_t)$ is increasing in z_t by A1, the payoff to deviation of the form $z_t < \widetilde{\tau}_t$ is increasing in z_t. Therefore, the supremum of the payoff to this kind of deviation is

$$\lim_{z_t \to \widetilde{\tau}_t} [w(z_t, \tilde{\tau}_t)(1-\delta)w(z_t, \tilde{\tau}_t) + (1-\delta)\sum_{s=t+1}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t)] = (1-\delta)\sum_{s=t}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t)$$

(b) If a country deviates to $z_t > \widetilde{\tau}_t$, it receives

$$g(z_t, \tilde{\tau}_t) = (1-\delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t) \tag{A.1}$$

So, it suffices to show that (A.1) has a global maximum z_t^* on (τ_t, ∞). If this is *not* the case, then there exists an increasing sequence $\{z^n\}$ with $\lim_{n\to\infty} z^n \to \infty$, for which $g(z^n, \tau_t)$ is monotonically increasing. But, for z^n high enough, the resulting consumption bundle, call it $\mathbf{x}(z^n, \tau_t)$, must be close to the autarchy allocation, and by the Inada conditions on utility, this will yield the consumer in the deviating country a lower utility than (for example) the bundle $\mathbf{x}(\tau_t, \tau_t)$ generated by not deviating. Contradiction. □

Proof of Lemma 2. (i) By definition, $z(\tau) = \max\{\zeta(\tau), \tau\}$. Moreover, as $\zeta(.)$ is decreasing in τ, it must be the case that there exists a $\overline{\tau}$ for which $\zeta(\tau) > \tau$, $\tau < \overline{\tau}$, $\zeta(\tau) < \tau$, $\tau > \overline{\tau}$.

(ii) We now prove that $\overline{\tau} < \hat{\tau}$. Suppose not; consider $\overline{\tau} = \hat{\tau}$ first. By the definition of (3.3) we must have $\zeta(\hat{\tau}) = \hat{\tau} = \arg\max_z \{w(\hat{\tau},\hat{\tau}) + \delta w(\hat{\tau},\hat{\tau})/(1-\delta)\}$. The first order condition requires that

$$w_1(\hat{\tau},\hat{\tau}) + \frac{\delta}{1-\delta}(w_1(\hat{\tau},\hat{\tau}) + w_2(\hat{\tau},\hat{\tau})) = 0$$

But by a standard argument, the myopic best response tariff $\hat{\tau}$ solves $w_1(\hat{\tau},\hat{\tau}) = 0$. By A2, we have that $w_1(\hat{\tau},\hat{\tau}) + w_2(\hat{\tau},\hat{\tau}) < 0$. Therefore, the first order condition cannot be satisfied at $\overline{\tau} = \hat{\tau}$; a contradiction. Then $\overline{\tau} > \hat{\tau}$ can also be ruled out because $w_1(\overline{\tau},\overline{\tau}) < 0$ for $\overline{\tau} > \hat{\tau}$.

Combining the fact that $z(\tau) = \max\{\zeta(\tau),\tau\}$ and the fact that there exists a unique $\overline{\tau}$ for which $\overline{\tau} = \zeta(\tau)$ with $\zeta(\tau)$ declining in τ, we see that $z(\tau) = \zeta(\tau)$ for $\tau < \overline{\tau}$ and $z(\tau) = \tau$ for $\tau \geq \overline{\tau}$.

(iii) To see that $z(0) > 0$, suppose to the contrary that $z(0) = 0$. Note that by the optimality of free trade, $w(0,0) > w(\tau,\tau)$, $\tau \neq 0$, which of course implies that

$$w_1(0,0) + w_2(0,0) = 0$$

Now, consider a small increase in z_t from 0, say Δ. Then, the effect of this change in z_t on the deviation payoff is

$$\Delta\left[(1-\delta)w_1(0,0) + \delta(w_1(0,0) + w_2(0,0))\right] = (1-\delta)\Delta w_1(0,0) > 0$$

where the last inequality follows from A1. □

Proof of Proposition 1. Suppose not. Suppose instead that after country i initially deviates in period t to $z(\tau_t)$ and country j matches the deviation in $t+1$ by also setting $z(\tau_t)$, there is an incentive for some country to deviate from $z(\tau_t)$ in some period $s > t+1$. By Lemma 2, the initial deviation has two characterizations: (i) if $\tau_t < \overline{\tau}$ then $z(\tau_t) = \zeta(\tau) \geq \overline{\tau} > \tau$; (ii) if $\tau_t \geq \overline{\tau}$ then $z(\tau_t) = \tau_t$. Each will be taken in turn.

(i) If $\tau_t < \overline{\tau}$ and so $z(\tau_t) = \zeta(\tau_t) \geq \overline{\tau} > \tau_t$ then, by Lemma 2 (ii), the best response to $z(\tau_t)$ must also be $z(\tau_t)$. To see why, observe that for $z(\tau_t) > \overline{\tau}$ the unconstrained best response given by (3.3) is $\zeta(\tau_t) < \overline{\tau}$ but that the constraint $z_t \geq \tau_t$ in (3.2) is binding. Therefore, $z(\tau_t)$ is a best response to itself; contradiction. Of course, $\overline{\tau}$ is self enforcing, so the contradiction is immediate.

(ii) If $\tau_t \geq \overline{\tau}$ then $z(\tau_t) = \tau_t \geq \overline{\tau}$ and so, again by Lemma 2 (ii), the best response to $z(\tau_t)$ must also be $z(\tau_t)$; contradiction. □

Proof of Proposition 2. Suppose to the contrary that WEC is not a subgame perfect punishment strategy; $z(\tau_t)$ is not a best response given the initial deviation $z(\tau_t)$. Henceforth we write $z(\tau_t)$ as z. Let τ' be a best response to z given suspension of GATT obligations next period, triggering $\hat{\tau}$ by both countries. Then choosing τ' in some period after initial deviation must bring about a higher payoff than adherence to WEC; that is,

$$(1-\delta)\, w(\tau', z) + \delta w(\hat{\tau}, \hat{\tau}) > (1-\delta)\, w(z, z) + \delta w(z, z)\,.$$

Note that, by Lemma 2(iii), $z > 0$ (requiring A1). Also note that the left and right hand sides of the inequality are consistent with one another because the initial deviation z is the same whether there is adherence to WEC or a suspension of GATT obligations.

Given that tariffs under suspension of GATT obligations are fixed at $\hat{\tau}$, we have that τ' is chosen to solve $w_1(\tau', z) = 0$. Using values under a suspension of GATT obligations in the appropriate first order condition of (3.1),

$$w_1(\tau', z) + \frac{\delta}{1-\delta}\left(w_1(\hat{\tau}, \hat{\tau}) + w_2(\hat{\tau}, \hat{\tau})\right) < 0,$$

because $w_1(\hat{\tau}, \hat{\tau}) + w_2(\hat{\tau}, \hat{\tau}) < 0$ by A2 and $\hat{\tau} > 0$. Now note that z is chosen so that

$$w_1(z, z) + \frac{\delta}{1-\delta}\left(w_1(z, z) + w_2(z, z)\right) = 0,$$

ie it must be the case that $w_1(z, z) > 0$ because $(w_1(z, z) + w_2(z, z)) < 0$ by A2 and $z > 0$. Indeed, given initial deviation at z, by construction payoffs are optimized by choosing z in response. So $(1-\delta)\, w(\tau', z) + \delta w(\hat{\tau}, \hat{\tau}) < (1-\delta)\, w(z, z) + \delta w(z, z)$; a contradiction. □

Proof of Proposition 3. Suppose to the contrary that $\tau_t = 0$ for some t. Then, at t, the incentive constraint is

$$(1-\delta)w(0,0) + \delta w(0,0) \geq (1-\delta)w(z(0), 0) + \delta w(z(0), z(0)) \qquad \text{(A.2)}$$

But recall that, by Lemma 2(iii) at the solution to (3.2) $z(0) > 0$. This implies

$$(1-\delta)w(z(0), 0) + \delta w(z(0), z(0)) > (1-\delta)w(0,0) + \delta w(0,0),$$

contradicting (3.5). □

Proof of Proposition 4. The only part that does not follow directly from Figure 1 is that $\tau^* = \overline{\tau}$. To prove this, it is sufficient to show that on the interval $[0, \overline{\tau}]$, the slope of α is greater than the slope of β in absolute value. This slope condition clearly rules out the case in Figure 1, where $\tau^* < \overline{\tau}$.[32] Now, the slope of β is

$$\beta'(\tau) = w_1(\tau, \tau) + w_2(\tau, \tau) \tag{A.3}$$

Moreover, from Lemma 2, the constraint $z \geq \tau$ is not binding on $[0, \overline{\tau}]$, so differentiating α and applying the envelope theorem gives:

$$\alpha'(\tau) = (1 - \delta) w_2(z, \tau) \tag{A.4}$$

Given $z \geq \tau$ in (A.4), we must have

$$w_2(z, \tau) - w_2(\tau, \tau) = \int_\tau^z [w_{12}] \, dx,$$

and from A3 we have $w_2(z, \tau) - w_2(\tau, \tau) < 0$, so

$$\alpha'(\tau) \leq (1 - \delta) w_2(\tau, \tau). \tag{A.5}$$

So, from (A.3), (A.5), the required condition is that

$$(1 - \delta) w_2(\tau, \tau) < w_1(\tau, \tau) + w_2(\tau, \tau)$$

Rearranging, this is

$$0 < w_1(\tau, \tau) + \delta w_2(\tau, \tau) \tag{A.6}$$

But, the FOC defining $\overline{\tau}$ is:

$$w_1(\overline{\tau}, \overline{\tau}) + \delta w_2(\overline{\tau}, \overline{\tau}) = 0 \tag{A.7}$$

As $\tau < \overline{\tau}$, from (A.7) we must have:

$$w_1(\tau, \tau) + \delta w_2(\tau, \tau) = -\int_\tau^{\overline{\tau}} [w_{11} + (1 + \delta) w_{12} + \delta w_{22}] dx \tag{A.8}$$

where the derivatives on the RHS of (A.8) are evaluated at (x, x). By A3, $w_{12} < 0$. By assumption, $w_{11}, w_{22} \leq 0$. So, (A.8) implies (A.6), as required.

[32] The case shown in Figure 2, where $\tau^* < \overline{\tau}$, requires that the slope of α must be less than that of β in absolute value somewhere in the interval $[\tau^*, \overline{\tau}]$.

The fact that $\tau^* = \overline{\tau} < \hat{\tau}$ follows from Lemma 2. □

Proof of Proposition 5. (a) Following the proof of Lockwood and Thomas (2002), Lemma 2.2, which uses assumptions analogous to A1-A3, the equilibrium conditions (3.6) can be shown to be equivalent to the following difference equation:

$$\alpha(\tau_{t+1}) = \frac{1}{\delta}\left[\alpha(\tau_t) - (1-\delta)\beta(\tau_t)\right],\ t = 1, .. \tag{A.9}$$

with initial condition $\tau_0 = \hat{\tau}$, plus the condition that the solution to (A.9) is bounded. To see this, note first that advancing the equality in (A.9) by one period (i.e. from t to $t+1$), multiplying the $t+1-$condition by δ and subtracting from the $t-$condition, we get:

$$\begin{aligned}(1-\delta)w(\tau_t,\tau_t) &= (1-\delta)w(z(\tau_t),\tau_t) + \delta w(z(\tau_t),(z(\tau_t)) \\ &-\delta\left[(1-\delta)w(z(\tau_{t+1}),\tau_{t+1}) + \delta w(z(\tau_{t+1}),(z(\tau_{t+1}))\right],\ t = 1, ..\end{aligned} \tag{A.10}$$

Using the definitions of α, β in (A.10) and rearranging, we get[33] (A.9).

(b) Now suppose that the path $\{\tau_t\}$ is in E and more efficient than the stationary path $\overline{\tau}$. Then, for some t, $\tau_t < \overline{\tau}$ (otherwise, $\tau_t \geq \overline{\tau}$, all t, so it cannot be more efficient). We now show that if $\tau_t < \overline{\tau}$, then $\tau_{t+1} < \tau_t$. For suppose not. Then, as α is decreasing in τ_t, we would have

$$\alpha(\tau_{t+1}) \leq \alpha(\tau_t) \tag{A.11}$$

Combining (A.9) and (A.11), we have

$$\frac{1}{\delta}\left[\alpha(\tau_t) - (1-\delta)\beta(\tau_t)\right] \leq \beta(\tau_t) \Longrightarrow \alpha(\tau_t) \leq \beta(\tau_t)$$

But as $\tau_t < \overline{\tau}$, $\alpha(\tau_t) > \beta(\tau_t)$, a contradiction. So, any solution of (A.9) is clearly a strictly decreasing sequence. There are then two possibilities. First, $\lim_{t\to\infty}\tau_t = \tau_\infty > \infty$. But then $\alpha(\tau_\infty) = \beta(\tau_\infty)$, contradicting the definition of $\overline{\tau} > \tau_\infty$ as the smallest root of $\alpha(\tau) = \beta(\tau)$. The other is $\lim_{t\to\infty}\tau_t = -\infty$. But this path cannot be more efficient than the stationary path, a contradiction. □

[33] The converse result can be obtained by solving (A.9) forward by substitution to get:

$$\alpha(\tau_t) = (1-\delta)(\beta(\tau_t) + \delta\beta(\tau_{t+1}) + ..\delta^n\beta(\tau_{t+n})) + \delta^{n-1}\alpha(\tau_{t+n+1})$$

So, as long as $\lim_{t\to\infty}\alpha(\tau_t) = 0$, (A.9) implies (3.6).

Proof of Proposition 6. Note that a tariff profile in which $\hat{\tau}$ is set in every period is an equilibrium path because the subgame in each period is a Nash equilibrium. But there is no incentive for either country to break a past binding by deviating upwards from such a path. So we restrict attention to a candidate equilibrium path for which $\widetilde{\tau}_t < \hat{\tau}$ at some t. Also, it may be the case that the optimal deviation from $\widetilde{\tau}_t$ entails $z(\tau_t) \leq \tau_{t-1}$, in which a deviant has no incentive to break a past binding. Therefore, we only need to consider situations where $z(\tau_t) > \tau_{t-1}$.

Normalize so that $t = 1$ is the first period in which $\widetilde{\tau}_t < \hat{\tau}$; let $\tau_0 = \hat{\tau}$ and $\widetilde{\tau}_1 = \widetilde{\tau} < \hat{\tau}$. We are interested in situations where a country has an incentive to break a past binding, that is, to set $\tau_t = \tau' > \tau_{t-1}$. Suppose to the contrary that no $\delta < 1$ can be found for which a deviant fails to break a past binding. Then in some period $t \geq 2$ the deviant, country i, must find it optimal to set a tariff $\tau_t = \tau' > \tau_{t-1}$ given that country j sets $\tau_t = \widetilde{\tau}_t$. (Note that we use $t \geq 2$ here as there have been no *past* bindings at period $t = 1$.) If the deviant sets $\tau' > \tau_{t-1}$ then this triggers $\hat{\tau}$ in all future periods, and this must yield a higher payoff than setting τ_{t-1} and facing WEC in all future periods:

$$(1-\delta)\, w(\tau', \widetilde{\tau}_t) + \delta w(\hat{\tau}, \hat{\tau}) > (1-\delta)\, w(\tau_{t-1}, \widetilde{\tau}_t) + \delta w(\tau_{t-1}, \tau_{t-1}).$$

By A1, A3, the largest possible gains from breaking a past binding occur for $\widetilde{\tau}_t = 0$. Now $\tau_{t-1} \leq \widetilde{\tau}$. So if we can show that there exists a $\delta < 1$ for which

$$(1-\delta)\, w(\tau', 0) + \delta w(\hat{\tau}, \hat{\tau}) < (1-\delta)\, w(\tau_{t-1}, \widetilde{\tau}_t) + \delta w(\widetilde{\tau}, \widetilde{\tau})$$

then we have established a contradiction. First note that, by A2, $w(\widetilde{\tau}, \widetilde{\tau}) \leq w(\tau_{t-1}, \tau_{t-1})$. But $\widetilde{\tau} < \hat{\tau}$ and so, again by A2, $w(\hat{\tau}, \hat{\tau}) < w(\widetilde{\tau}, \widetilde{\tau})$. Therefore, as the inequality holds strictly at the $\delta = 1$ limit, it must hold for a range of δ less than 1. Also, as it holds for $w(\tau', 0)$, it must hold for all $\widetilde{\tau}_t > 0$ under which the gains to breaking past bindings are smaller. This establishes a contradiction. □

Proof of Lemma 3. By Lemma 2 there exists a unique solution $\overline{\tau}$ for any $\delta \in (0, 1)$. By Lemma 2(ii), $z(\tau_t) \geq \overline{\tau}$ for any $\tau_t \leq \overline{\tau}$ and by assumption $\overline{\tau} \geq \tau_{t-1}$. Then by definition, $\chi(\tau_t, \tau_{t-1}) = \tau_{t-1}$. □

Proof of Lemma 4. The proof is by induction. By assumption $\tau_t < \tau_{t-1} < \overline{\tau}$.

From (4.1) and Lemma 3, an efficient path with $\tau_t \leq \overline{\tau}$, $t \geq 1$ must satisfy

$$\begin{aligned}(1-\delta)(w(\tau_t,\tau_t)+\delta w(\tau_{t+1},\tau_{t+1})+...\quad) &= \\ (1-\delta)w(\tau_{t-1},\tau_t)+\delta w(\tau_{t-1},\tau_{t-1}),\ t &= 1,...\end{aligned} \tag{A.12}$$

Advancing (A.12) one period, multiplying both sides by δ, subtracting from (A.12), and dividing the result by $1-\delta$, we get:

$$w(\tau_t,\tau_t) = w(\tau_{t-1},\tau_t) + \frac{\delta}{1-\delta}w(\tau_{t-1},\tau_{t-1}) - \delta\left[w(\tau_t,\tau_{t+1}) + \frac{\delta}{1-\delta}w(\tau_t,\tau_t)\right] \tag{A.13}$$

which is a second-order difference equation[34] in τ_t. This can be seen more clearly by rearranging (A.13) to get:

$$w(\tau_t,\tau_{t+1}) = \frac{1}{\delta}\left[w(\tau_{t-1},\tau_t) - w(\tau_t,\tau_t)\right] + \frac{w(\tau_{t-1},\tau_{t-1})}{1-\delta} - \frac{\delta w(\tau_t,\tau_t)}{1-\delta},\ t>1.$$

Rewriting (4.2), we get:

$$\delta\left[w(\tau_t,\tau_{t+1}) - w(\tau_t,\tau_t)\right] = w(\tau_{t-1},\tau_t) + \frac{\delta w(\tau_{t-1},\tau_{t-1})}{1-\delta} - \left[w(\tau_t,\tau_t) + \frac{\delta w(\tau_t,\tau_t)}{1-\delta}\right]$$

By Proposition 6 and Lemma 3 we know that for $\delta \in (\underline{\delta},1)$ and for $\tau_t < \tau_{t-1} < \overline{\tau}$ the solution to the constrained optimization problem

$$\max_{\tau_t<z_t<\tau_{t-1}}\left\{w(z_t,\tau_t) + \frac{\delta w(z_t,z_t)}{1-\delta}\right\}$$

is τ_{t-1}. So we can write

$$\begin{aligned}& w(\tau_{t-1},\tau_t) + \frac{\delta w(\tau_{t-1},\tau_{t-1})}{1-\delta} - \left[w(\tau_t,\tau_t) + \frac{\delta w(\tau_t,\tau_t)}{1-\delta}\right] \\ = & \max_{\tau_t<z_t<\tau_{t-1}}\left\{w(z_t,\tau_t) + \frac{\delta w(z_t,z_t)}{1-\delta}\right\} - \left[w(\tau_t,\tau_t) + \frac{\delta w(\tau_t,\tau_t)}{1-\delta}\right] \\ > & \ 0.\end{aligned}$$

Therefore $w(\tau_t,\tau_{t+1}) > w(\tau_t,\tau_t)$. But then, by A1, $\tau_{t+1} < \tau_t$ as required. □

Proof of Proposition 7. In order for $\chi(\tau_t,\tau_{t-1}) = \tau_{t-1}$ to be the optimal deviation, by Proposition 6, we require that $\delta \in (\underline{\delta},1)$. Now, rewrite (4.3) as a function of τ_1 :

$$\begin{aligned}f(\tau_1) &= (1-\delta)w(\chi(\tau_1,\hat{\tau}),\tau_1) + \delta w(\chi(\tau_1,\hat{\tau}),\chi(\tau_1,\hat{\tau})) \\ &\quad -(1-\delta)(w(\tau_1,\tau_1) + \delta w(\tau_2(\hat{\tau},\tau_1),\tau_2(\hat{\tau},\tau_1)) + ...)\end{aligned}$$

[34] This is an unusual difference equation in that it has a continuum of stationary solutions i.e. setting $\tau_{t-1} = \tau_t = \tau_{t+1}$ always solves (2.1).

Note that by the definition of $\overline{\tau}$ (see Lemma 2)

$$(1-\delta)w(\chi(\overline{\tau},\hat{\tau}),\overline{\tau})+\delta w(\chi(\overline{\tau},\hat{\tau}),\chi(\overline{\tau},\hat{\tau}))=w(\overline{\tau},\overline{\tau})$$

Moreover, $\tau_t(\hat{\tau},\overline{\tau})<\overline{\tau}$, all t by Lemma 4. So, if $\tau_1=\overline{\tau}$ then (4.1) is slack i.e.

$$\begin{aligned}&(1-\delta)(w(\overline{\tau},\overline{\tau})+\delta w(\tau_2(\hat{\tau},\overline{\tau}),\tau_2(\hat{\tau},\overline{\tau}))+...)\\ > \ & w(\overline{\tau},\overline{\tau})=(1-\delta)w(\chi(\overline{\tau},\hat{\tau}),\overline{\tau})+\delta w(\chi(\overline{\tau},\hat{\tau}),\chi(\overline{\tau},\hat{\tau}))\end{aligned}$$

where the inequality follows by A2. So, we have shown that $f(\overline{\tau})<0$.

Next, if $\tau_1=\varepsilon$, we have

$$(1-\delta)w(\chi(\varepsilon,\hat{\tau}),\varepsilon)+\delta w(\chi(\varepsilon,\hat{\tau}),\chi(\varepsilon,\hat{\tau}))=\max_{\varepsilon\leq z\leq\hat{\tau}}(1-\delta)w(z,\varepsilon)+\delta w(z,z)>w(\varepsilon,\varepsilon)$$

for ε small enough: the inequality is strict by Lemma 2 above, as for ε small enough, $z(\varepsilon)>\varepsilon$. Moreover, from Lemma 4, for ε small enough,

$$(1-\delta)(w(\varepsilon,\varepsilon)+\delta w(\tau_2(\hat{\tau},\varepsilon),\tau_2(\hat{\tau},\varepsilon))+...)\simeq w(\varepsilon,\varepsilon)$$

So, it is possible to choose ε small enough so that

$$(1-\delta)(w(\varepsilon,\varepsilon)+\delta w(\tau_2(\hat{\tau},\varepsilon),\tau_2(\hat{\tau},\varepsilon))+...)<(1-\delta)w(\chi(\varepsilon,\hat{\tau}),\varepsilon)+\delta w(\chi(\varepsilon,\hat{\tau}),\chi(\varepsilon,\hat{\tau}))$$

i.e. $f(\varepsilon)>0$. Now, by inspection, $f(.)$ is continuous in τ_1 as χ and τ_t are continuous in τ_1. So, there exists at least one value of τ_1 for which $f(\tau_1)=0$, and so there exists a smallest such value. □

A.2. An Example: Quasi-linear Preferences

We work through the analysis of the paper for a simple endowments model where agents have quasi-linear preferences. Each country $i\in\{1,2\}$ has an endowment (normalized to unity) of good i (or is endowed with a factor of production that can produce 1 unit of good i). We denote by x_j^i the consumption of good j in country i. The preferences of the representative consumer in country i are of the following form:

$$u^i=x_i^i+\frac{\sigma}{\sigma-1}{x_j^i}^{\frac{\sigma-1}{\sigma}},\ i=1,2 \qquad \text{(A.14)}$$

with $\sigma > 1$, and where x_j^i is consumption of good j. σ measures the elasticity if substitution between different "varieties" of imported goods. Utility is maximized subject to the budget constraint

$$p_i x_i^i + p_j(1+\tau^i)x_j^i = p_i + R_i \tag{A.15}$$

where p_j, and R_i are respectively: the world price of good j and tariff revenue in country i which, as is usually assumed, is returned to the consumer in a lump-sum. Note that whilst in the text above we referred to country i's tariff in period t as τ_t^i here time subscripts are dropped. The optimization problem gives demands for the two goods;

$$x_j^i = \left[\frac{p_j(1+\tau^i)}{p_i}\right]^{-\sigma}, \ j \neq i \tag{A.16}$$

$$x_i^i = 1 + \frac{R_i}{p_i} - \frac{p_j(1+\tau^i)x_j^i}{p_i} = 1 + \frac{R_i}{p_i} - \left[\frac{p_j(1+\tau^i)}{p_i}\right]^{1-\sigma} \tag{A.17}$$

where the demand for good i, x_i^i is determined residually via the budget constraint.

Indirect utility for the representative household in i is therefore derived by substituting (A.16) ,(A.17), back into (A.14) to get

$$v^i = \frac{1}{\sigma-1}\left[\frac{p_j(1+\tau^i)}{p_i}\right]^{1-\sigma} + \frac{R_i}{p_i} \tag{A.18}$$

Also, tariff revenue is

$$R_i = p_j \tau^i x_j^i = \frac{p_j \tau^i}{p_i}\left[\frac{p_j(1+\tau^i)}{p_i}\right]^{-\sigma} \tag{A.19}$$

We substitute (A.19) into (A.18) to get:

$$v^i = \frac{1}{\sigma-1}\left[\frac{p_j(1+\tau^i)}{p_i}\right]^{1-\sigma} + \frac{p_j \tau^i}{p_i}\left[\frac{p_j(1+\tau^i)}{p_i}\right]^{-\sigma} \tag{A.20}$$

We may choose p_i as the numeraire write (A.20) as

$$v(\tau, p) = \frac{1}{\sigma-1}\left[p(1+\tau)\right]^{1-\sigma} + p\tau\left[p(1+\tau)\right]^{-\sigma} \tag{A.21}$$

Finally, we need to calculate how the (reciprocal of) terms of trade for country i, p, changes with τ', τ. Evaluating (A.16) ,(A.17) at $\tau^j = \tau'$ $\tau^i = \tau$, $p_j = p$, $p_i = 1$, we get;

$$x_i^i = 1 + p\tau\left[p(1+\tau)\right]^{-\sigma} - \left[p(1+\tau)\right]^{1-\sigma} \tag{A.22}$$

$$x_i^j = \left[\frac{(1+\tau')}{p}\right]^{-\sigma} \tag{A.23}$$

So, substituting (A.22),(A.23) into the market-clearing condition for good i, namely that supply of unity equals the sum of country demands ($1 = \sum_{i\in\{1,2\}} x_i^j$), we have

$$p\tau\left[p(1+\tau)\right]^{-\sigma} - \left[p(1+\tau)\right]^{1-\sigma} + \left[\frac{(1+\tau')}{p}\right]^{-\sigma} = 0 \tag{A.24}$$

Solving (A.24) for p, we get:

$$p(\tau,\tau') = \left(\frac{1+\tau}{1+\tau'}\right)^{\sigma/(1-2\sigma)}$$

Note that as $\sigma > 0.5$ by assumption, $p_\tau < 0$ i.e. an increase in $i's$ tariff always improves $i's$ terms of trade. So, we may write country $i's$ indirect utility as

$$w(\tau,\tau') \equiv v(p(\tau,\tau'),\tau) = \frac{1}{\sigma-1}\left[p(1+\tau)\right]^{1-\sigma} + p\tau\left[p(1+\tau)\right]^{-\sigma}$$

So, a (symmetric) Nash equilibrium in tariffs is a $\hat{\tau}$ such that $v(\hat{\tau}, p(\hat{\tau},\hat{\tau})) \geq v(\tau, p(\tau,\hat{\tau}))$, all $\tau \neq \hat{\tau}$.

As v is continuously differentiable, we can characterize $\hat{\tau}$ as the solution to

$$v_\tau(\hat{\tau}, p(\hat{\tau},\hat{\tau})) + v_p(\hat{\tau}, p(\hat{\tau},\hat{\tau}))p_\tau(\hat{\tau},\hat{\tau}) = 0 \tag{A.25}$$

where v_τ, v_p denote partial derivatives of v. Now,

$$\begin{aligned} v_\tau(\tau,p) &= -\sigma\tau p^{1-\sigma}(1+\tau)^{-\sigma-1} \\ v_p(\tau,p) &= -p^{-\sigma}(1+\tau)^{1-\sigma} + (1-\sigma)p^{-\sigma}\tau(1+\tau)^{-\sigma} \\ p_\tau &= \frac{\sigma}{1-2\sigma}\left(\frac{1+\tau}{1+\tau'}\right)^{(\sigma/(1-2\sigma))-1}\frac{1}{1+\tau'} \end{aligned} \tag{A.26}$$

So, using (A.26) and the fact that $p(\hat{\tau},\hat{\tau}) = 1$, we have from (A.25) that

$$-\sigma\hat{\tau}(1+\hat{\tau})^{-\sigma-1} + \left[-(1+\hat{\tau})^{1-\sigma} + (1-\sigma)\hat{\tau}(1+\hat{\tau})^{-\sigma}\right]\frac{\sigma}{1-2\sigma}\frac{1}{1+\hat{\tau}} = 0$$

Eliminating common terms, we get

$$-\hat{\tau} + \left[-(1+\hat{\tau}) + (1-\sigma)\hat{\tau}\right]\frac{1}{1-2\sigma} = 0$$

Solving, we get

$$\hat{\tau} = \frac{1}{\sigma-1}$$

for the optimal tariff. Recall that $\sigma > 1$, so $\hat{\tau}$ is defined and positive.

Now we have $\hat{\tau}$, we can check that A1, A2 and A3 hold for tariffs set on the interval $[0, \hat{\tau}]$

Substituting for $p(\tau, \tau')$, we can write the payoff function as follows:

$$w(\tau, \tau') = \left(\frac{(1+\tau)^{1-\sigma}}{\sigma - 1} + \tau (1+\tau)^{-\sigma} \right) \left(\frac{1+\tau}{1+\tau'} \right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

We can use this expression to verify that A1, A2 and A3 hold. Take A1 first:

$$w_1(\tau, \tau') = \frac{\sigma (1+\tau)^{-1-\sigma} (1 - (\sigma - 1)\tau)}{2\sigma - 1} \left(\frac{1+\tau}{1+\tau'} \right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

The sign of this expression depends on the term in brackets $(1 - (\sigma - 1)\tau)$. If $\tau = \hat{\tau} = 1/(\sigma - 1)$ and $(1 - (\sigma - 1)\tau) = 0$ so $w_1(\tau, \tau') = 0$. If $\tau < \hat{\tau}$ then $(1 - (\sigma - 1)\tau) > 0$ and so $w_1(\tau, \tau') > 0$ as required.

$$w_2(\tau, \tau') = -\frac{\sigma (1+\tau)^{-1-\sigma} (1 + \sigma\tau)}{2\sigma - 1} \left(\frac{1+\tau}{1+\tau'} \right)^{\left(1-\sigma-\sigma^2\right)/(1-2\sigma)} < 0 \text{ for all } \tau, \tau' \geq 0.$$

Now A2:

$$w_1(\tau, \tau') + w_2(\tau, \tau') = -\frac{\sigma (1+\tau)^{-2-\sigma} (\sigma\tau (2 + \tau + \tau') - (1+\tau)\tau')}{2\sigma - 1} \left(\frac{1+\tau}{1+\tau'} \right)^{\left(1-\sigma-\sigma^2\right)/(1-2\sigma)}$$

Now the sign of this expression depends on the term in brackets $(\sigma\tau (2 + \tau + \tau') - (1+\tau)\tau')$. It is easy to see that when $\tau = \tau' = 0$ we have $(\sigma\tau (2 + \tau + \tau') - (1+\tau)\tau') = 0$ and therefore $w_1(\tau, \tau') + w_2(\tau, \tau') = 0$. This is necessary for free trade to maximize efficiency. Moreover, by inspection $(\sigma\tau (2 + \tau + \tau') - (1+\tau)\tau') > 0$ for all $\tau, \tau' \in (0, \hat{\tau})$, $\sigma > 1$, so $w_1(\tau, \tau') + w_2(\tau, \tau') < 0$ as required. Finally, regarding A3:

$$w_{12}(\tau, \tau') = -\frac{(\sigma - 1)\sigma^2 (1+\tau)^{-2-\sigma} (1 - (\sigma - 1)\tau)}{(2\sigma - 1)^2} \left(\frac{1+\tau}{1+\tau'} \right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

So $w_{12}(\tau, \tau') < 0$ because $(1 - (\sigma - 1)\tau) > 0$ for $\tau, \tau' \in (0, \hat{\tau})$ as required.

Now we want to characterize the constrained deviation, using it to derive $\overline{\tau}$. Setting this first order condition equal to zero, we have

$$w_1(z(\tau), \tau) + \frac{\delta}{1-\delta} (w_1(z(\tau), z(\tau)) + w_2(z(\tau), z(\tau))) = 0.$$

We can write (A.14) as follows

$$w(z(\tau), \tau) = \left(\frac{1+z(\tau)}{1+\tau}\right)^{\sigma(1-\sigma)/(1-2\sigma)} \gamma(z(\tau)),$$

where $\gamma(z(\tau)) = \frac{(1+z(\tau))^{1-\sigma}}{\sigma-1} + z(\tau)(1+z(\tau))^{-\sigma}$, so $\gamma'(z(\tau)) = -\sigma z(\tau)(1+z(\tau))^{-1-\sigma}$ .Then

$$w_1(z(\tau), \tau) = \frac{\frac{\sigma(1-\sigma)}{1-2\sigma} w(z(\tau), \tau)}{(1+z(\tau))} + \left(\frac{1+z(\tau)}{1+\tau}\right)^{\frac{\sigma(1-\sigma)}{1-2\sigma}} \gamma'(z(\tau)),$$

and

$$w_2(z(\tau), \tau) = -\frac{\frac{\sigma(1-\sigma)}{1-2\sigma} w(z(\tau), \tau)}{(1+\tau)}$$

It is then straightforward to see that the first order condition can be rewritten $(1-\delta) w_1(z(\tau), \tau) + \delta\gamma'(z(\tau)) = 0$.Setting $z(\tau) = \tau = \overline{\tau}$ in the first order condition, we get

$$(1-\delta)\frac{\sigma(\sigma-1)}{2\sigma-1}\frac{\gamma(\overline{\tau})}{1+\tau^*} + \gamma'(\overline{\tau}) = 0$$

Substituting for $\gamma(\overline{\tau})$ and $\gamma'(\overline{\tau})$ and simplifying, the equation becomes

$$\frac{\sigma(1+\overline{\tau})^{-1-\sigma}(1-\delta+(1-\sigma(1+\delta))\overline{\tau})}{2\sigma-1} = 0$$

Solving, the only admissible root[35] is

$$\overline{\tau} = \frac{1-\delta}{\sigma(1+\delta)-1}.$$

[35] The root $\tau = -1$ also solves this expression.

References

[1] Bagwell, K. and R. Staiger (1990); "A Theory of Managed Trade." *American Economic Review*, 80, 779-795.

[2] Bagwell, K. and R. Staiger (1996); "Reciprocal Trade Liberalization." NBER Working Paper No. 5488.

[3] Bagwell, K. and R. Staiger (1997); "Reciprocity, Non-discrimination, and Preferential Agreements in the Multilateral Trading System." NBER Working Paper No. 5932.

[4] Bagwell, K. and R. Staiger (1999); "An Economic Theory of the GATT." *American Economic Review*, 89(1) (March): 215-48.

[5] Bagwell, K. and R. Staiger (2002); *The Economics of the World Trading System*, MIT Press, Cambridge, Massachusetts.

[6] Blume, A. "Intraplay Communication in Repeated Games." *Games and Economic Behavior*, 6, 181-211.

[7] Bond, E.B. and J. H. Park (2002); "Gradualism in Trade Agreements with Asymmetric Countries." *Review of Economic Studies*, 69(2): 379-406.

[8] Bond, E.B., R. Riezman and C. Syropoulos (2003); "A Strategic and Welfare Theoretic Analysis of Free Trade Areas." *Journal of International Economics*, forthcoming.

[9] Bond, E.B. and C. Syropoulos (1996); "Trading Blocs and the Sustainability of Interregional Cooperation." Chapter in M. Canzoneri, W. Ethier and V. Grilli (eds), *The New Transatlantic Economy*, pp. 118-141. Cambridge University Press, Cambridge, UK.

[10] Bond, E.B., C. Syropoulos and A. Winters (2001); "Deepening of regional integration and multilateral trade agreements." *Journal of International Economics*, 53: 335-361.

[11] Chisik, R. (2003); "Gradualism in Free Trade Agreements: A Theoretical Justification." *Journal of International Economics*, forthcoming.

[12] Dam, K.W. (1970); *The GATT: Law and International Economic Organization.* Chicago University Press, Chicago.

[13] Devereux, M. (1997); "Growth, specialization and trade liberalization." *International Economic Review*, 38(3): 565-585.

[14] Dixit, A. (1987); "Strategic Aspects of Trade Policy." Chapter 9 in Bewley, T. F. *Advances in Economic Theory; Fifth World Congress*, Cambridge University Press, Cambridge.

[15] Farrell, J. and E. Maskin (1989); "Renegotiation in Repeated Games." *Games and Economic Behavior*, 1, 327-360.

[16] Friedman, J. (1971); "A Noncooperative Equilibrium for Supergames." *Review of Economic Studies* 38: 1-12.

[17] Fudenberg, D. and E. Maskin (1986); "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica*, 54: 533-556.

[18] Fudenberg, D. and J. Tirole (1991); *Game Theory*, MIT Press, Cambridge Mass.

[19] Furusawa, T. and L.-C. Lai (1999); "Adjustment costs and gradual trade liberalization." *Journal of International Economics*, 49: 333-361.

[20] GATT (1962); "Report of the Panel on Uruguayan Recourse to Article XXIII." Available online at http://www.wto.org.

[21] GATT (1994); "Uruguay Round: Final Act," Marrakesh, 15 April. Photocopy.

[22] Grossman, G. and E. Helpman (1995) "Trade Wars and Trade Talks." *Journal of Political Economy*, 103(4): 675-708.

[23] Horwell, D.J. (1966); "Optimum Tariffs and Tariff Policy." *Review of Economic Studies*, 33, 147-158.

[24] Jackson, J. (1989); *The World Trading System.* MIT Press, Cambridge Massachusetts.

[25] Johnson, H. G. (1953-54); "Optimum Tariffs and Retaliation." *Review of Economic Studies*, 21, 142-153.

[26] Lockwood, B. and J. Thomas (2002); "Gradualism and Irreversibility." *Review of Economic Studies*, 69(2): 339-356.

[27] Lockwood, B. and B. Zissimos (2001); "Trade Agreements with Limited Punishments." Centre for the Study of Globalisation and Regionalisation Discussion Paper 87/02, University of Warwick.

[28] Lockwood, B. and K.Y. Wong (2000); "Specific and Ad Valorem Tariffs Are Not Equivalent in Trade Wars." *Journal of International Economics*, 52, 183-195.

[29] Mayer, W. (1981); "Theoretical Considerations of Negotiated Tariff Adjustments." *Oxford Economic Papers*, 33, 135-153.

[30] McCutcheon, B. (1997); "Do Meetings in Smoke-Filled Rooms Facilitate Collusion?" *Journal of Political Economy*, 105, 330-350.

[31] McLaren, J. (1997); "Size, Sunk Costs and Judge Bowker's Objection to Free Trade." *American Economic Review*, 87, 400-420.

[32] Scherer, F.M. (1980); *Industrial Market Structure and Economic Performance*, second edition, Houghton Mifflin.

[33] Staiger, R. (1995); "A theory of gradual trade liberalization." In Levinsohn, J., A.V. Deardorff and R.M. Stern (eds.) *New Directions in Trade Theory*, University of Michigan Press, Ann Arbor, MI.

[34] Syropoulos, C. (2002); "Optimum Tariffs and Retaliation Revisited: How Country Size Matters." *Review of Economic Studies*, 69(3): 707-727.

[35] van Damme, E. (1989); "Renegotiation-Proof Equilibria in Repeated Prisoners' Dilemma." *Journal of Economic Theory* 47, 206-217.

[36] Whalley, J. and C. Hamilton (1996); *The Trading System After The Uruguay Round*, Institute for International Economics, Washington, DC.

Figure 1

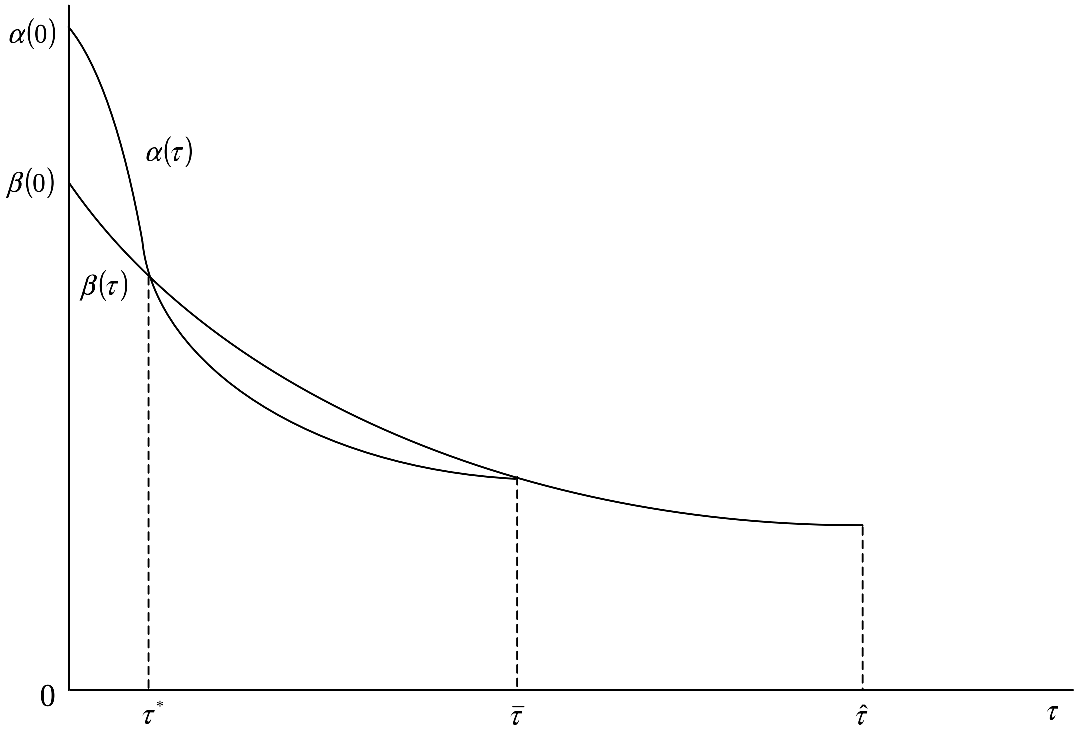

Figure 2