Trade Agreements with Limited Punishments[1]

Ben Lockwood and Ben Zissimos
CSGR and Department of Economics, University of Warwick
CSGR Working Paper No 87/02
First version: April 2001
This version: January 2002

Abstract

This paper shows that free trade can never be achieved when punishment for deviation from a trade agreement is limited to 'a withdrawal of equivalent concessions'. This is where retaliation is not allowed to entail higher tariffs than those set by the initial deviant, and is the most severe form of punishment allowed under WTO rules. If, in addition, deviations from agreements are also limited in some way, then efficient self-enforcing tariff reductions must be gradual.

Keywords: Nash tariffs, Free Trade, Gradualism, Trade agreement, World Trade Organisation (WTO)

JEL classification: FO2, F13, F15, C73

Address for correspondence
Professor Ben Lockwood & Ben Zissimos
CSGR
University of Warwick
Coventry
CV4 7AL, UK
Email: B.lockwood@warwick.ac.uk / b.c.zissimos@warwick.ac.uk

[1] We would like to thank Ray Riezman, John Whalley and Kar-yiu Wong for helpful comments and conversations about this paper. We are also grateful for comments on an earlier draft from seminar participants at the Spring 2001 Midwest International Meeting, University of Wisconsin, and the WTO and World Trade II Meeting, University of Washington. Support from the ESRC's Centre for the Study of Globalisation and Regionalisation, University of Warwick, is gratefully acknowledged.

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. Since the General Agreement of Tariffs and Trade (GATT) was drawn up after the war, tariffs have fallen from an average of 50 percent, to around 5 percent today. Second, tariffs have been cut only gradually in successive rounds of negotiations under the GATT (now the World Trade Organization, or WTO). Neither of these two facts sits well with the simple 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 of these two puzzles by modelling the rules imposed on trade negotiations by the WTO. In particular, we focus on the implications for the liberalization process of the WTO rule on the *withdrawal of equivalent concessions* (WEC) as set out in Article XXVIII of the GATT charter. Suppose that a deviant country fails to implement some agreed market access measure, whilst all other parties to the agreement proceed to do so. When the failure is discovered, under WTO rules trade partners 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 WTO. In terms of the applied game theory literature, WEC imposes *partial irreversibility* on punishments in this game. This is new, in that only complete irreversibility has been analyzed in the past (Lockwood and Thomas 2002).

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 our 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. 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.

Our second main result concerns the gradualism of trade liberalization. Specifically, 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. Article 2 of GATT (1994) in the Charter of the WTO specifies that a schedule of commitments be maintained. Results of tariff negotiations are recorded as scheduled commitments in the form of tariff bindings; a permanent and irrevocable commitment that tariffs will *not rise above bound levels* for the product in question. If tariffs are raised above bound levels, then we assume that this incurs a *loss of political good will*. Moreover, we suppose that the loss of political good will is so costly that it is never incurred in equilibrium. This implies that the optimal deviation is simply not to cut tariffs from the previous period's level (but not to raise them either). In this situation, because punishment is limited, current tariff cuts can only be made self enforcing by the promise of future tariff reductions. Moreover, if deviation can at worst entail not raising tariffs, then it is always possible to promise liberalization over a number of future periods that would more than compensate. This is gradualism in other words.

The paper builds on a substantial literature going back to Johnson (1953-54).[2] 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).[3] As remarked above, this trigger strategy approach 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 agreements are gradual. The

[2] 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.

[3] 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 (2001) 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.

general idea is that initially, full liberalization cannot be self-enforcing, as 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 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. In Devereux (1997), there is dynamic learning-by-doing in the export sector. In Furusawa and Lai (1999), there are linear[4] adjustment costs incurred when labor moves between sectors. Bond and Park (2000) consider gradualism in a framework where countries are asymmetric.

Finally, Bagwell and Staiger's (1999) work relates closely to our own, in that they too model specifically the GATT/WTO institutional framework. However, their focus is different. First, they make the very important point that the *only* thing that matters in a trade agreement is the terms of trade externality. This point is made very forcefully by constructing a model that is broader than ours in that it allows for a wider set of political variables to be present. And wider aspects of the GATT institutional framework than just the withdrawal of equivalent concessions are also examined in their work. But our model of withdrawal of equivalent concessions is built around a dynamic game, which theirs is not, and this enables us to bring out some implications of the institutional framework that they do not.[5] The theory of repeated games has also been used by Bond, Syropoulos and Winters (2001) to study trade block formation, where a *preferential* trade agreement is supported by the credible threat of punishment.

Our 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, showing that irreversibility on the side both of the initial deviant and the punisher are sufficient for gradualism. As pointed out above, in the first part of this present paper we assume (partial) irreversibility of the strategic instrument - here tariffs -

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

[5] The differences between Bagwell and Staiger's analysis and ours are discussed further in the Conclusions.

on the side of the punisher, but with the initial deviation itself unrestricted. We then show explicitly that gradualism cannot result. Only when there is a degree of irreversibility on both sides does gradualism arise. In this sense, the present paper extends Lockwood and Thomas (2002).

The paper proceeds as follows. The next section sets up the basic analytical framework, defines formally the tariff reduction game and a withdrawal of equivalent concessions. Section 3 then defines symmetric equilibrium tariff paths and examines their properties under a withdrawal of equivalent concessions. It is here that we show how trade liberalization is achieved in this framework but that free trade cannot be reached. Section 4 then examines the circumstances under which gradual trade liberalization can take place, presenting computed equilibrium tariff reduction paths for various parameterizations of a quasi-linear example. Section 5 concludes.

2. The Tariff Reduction Game

2.1. Tariffs and Welfare

We work with a simple and standard model of international trade. There are n countries $i \in N$ and the same number of goods. Each country i 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 over $\mathbf{x}^i = (x_j^i)_{j \in N}$ are then[6]

$$u^i(\mathbf{x}^i) = u(x_i^i, \varphi(\mathbf{x}^{-i})) \tag{2.1}$$

where $\mathbf{x}^{-i} = (x_1^i, ..x_{i-1}^i, x_{i+1}^i, ..x_n^i)$. Also, we assume that in equilibrium, some quantity of imported goods will be consumed i.e. we make the Inada-type assumption that $\lim_{\mathbf{x} \to 0} \partial u(x_i^i, \varphi(\mathbf{x}^{-i}))/\partial x_j^i = +\infty$, $j \neq i$. An example of this form is the quasi-linear utility function:

$$u^i = x_i^i + \frac{\sigma}{\sigma - 1} \sum_{j \neq i} (x_j^i)^{\frac{\sigma-1}{\sigma}}, \; i = 1, ..n \tag{2.2}$$

[6] We adopt the usual convention that bold characters denote vectors, and non-bold chracters denote scalars.

with $\sigma > 1$, and where σ measures the elasticity if substitution between different "varieties" of imported goods.

The consumer in country i faces a budget constraint

$$\sum_{j=1} p_j(1+\tau_j^i)x_j^i = p_i + R_i \tag{2.3}$$

where p_j, τ_j^i, R_i are respectively: the world price of good j, the tariff set by country i on good j, and tariff revenue in country i which, as is usually assumed, is returned to the consumer in a lump-sum. Without loss of generality, we set $\tau_i^i = 0$; also note that $-1 < \tau_j^i < \infty$.

Within a period, $t = 1, 2, \dots$, the order of events is as follows. First, each country i simultaneously chooses an import tariff vector $\boldsymbol{\tau}^i = (\tau_j^i)_{j\in N}$. Then, given world prices $\mathsf{p} = (p_j)_{j\in N}$, and $\boldsymbol{\tau}^i$, the consumer in country $i \in N$ chooses x^i to maximize u_i subject to the budget constraint, which yields the usual indirect utility function $v^i = v^i(\mathsf{p}, \boldsymbol{\tau}^i, R_i)$ and excess demands. Then, conditional on $\boldsymbol{\tau} = (\boldsymbol{\tau}^1, ..\boldsymbol{\tau}^n)$, markets clear and world prices p for the goods are determined.[7] These world prices will of course depend on tariffs i.e. $\mathsf{p} = \mathsf{p}(\boldsymbol{\tau})$, and so will tariff revenues i.e. $R_i = \sum_{j=1}^n p_j(\boldsymbol{\tau})\, \tau_j^i x_j^i (\mathsf{p}(\boldsymbol{\tau}))$. We assume that equilibrium prices are unique, given tariffs, so the mapping $\mathsf{p}(.)$ is one-to-one. It is also assumed that technology (embodied in u^i or v^i) is identical across countries.

So, we can write equilibrium welfare of country i, v^i, as a function of $\boldsymbol{\tau} = (\boldsymbol{\tau}^1, ..\boldsymbol{\tau}^n)$ only i.e. $v^i = v^i(\boldsymbol{\tau}^1, ..\boldsymbol{\tau}^n) \equiv v^i(\mathsf{p}(\boldsymbol{\tau}), \boldsymbol{\tau}^i, R_i(\boldsymbol{\tau}))$. Now we can define a *Nash equilibrium in tariffs* in the usual way as a $\hat{\boldsymbol{\tau}}$ such that $v^i(\hat{\boldsymbol{\tau}}^i, \hat{\boldsymbol{\tau}}^{-i}) \geq v^i(\boldsymbol{\tau}^i, \hat{\boldsymbol{\tau}}^{-i})$, all $\boldsymbol{\tau}^i \in (-1, \infty)^n$, all $i \in N$. We will focus on Nash equilibria where (i) all countries set *common tariffs* i.e. $\hat{\tau}_j^i = \hat{\tau}^i$, all $i \in N$; (ii) all these common tariffs are *equal* $\hat{\tau}^i = \hat{\tau}$, all $i \in N$. Such equilibria exist for the special cases that we consider below, due to the symmetry of the model[8].

[7] 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. This technical detail, and others, are dealt with in Section 3 below.

[8] More generally, it is possible to show that if all $j \neq i$ set the same common tariff, the unique best response of i is to set the same tariff on imports on all countries i.e. a common tariff.

2.2. The Tariff Reduction Game

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. $\tau^i_j = 0$, if the tariff reduction plan must be *self-enforcing* i.e. the outcome of a subgame-perfect equilibrium. It is convenient to impose the constraint that the cooperative tariff reductions have the same structure as does the Nash equilibrium i.e. each country sets a common tariff, τ^i. In this case, we may write country welfare as a function of common tariffs only i.e. $v^i = v^i(\tau^i, \tau^{-i})$. The following result establishes that, furthermore, countries have symmetric preferences over (common) tariffs.[9]

Proposition 1. $v^i = v(\tau^i, \tau^{-i})$, *and if* $\pi(\tau^{-i})$ *is any permutation of* τ^{-i}, *then* $v(\tau^i, \tau^{-i}) \equiv v(\tau^i, \pi(\tau^{-i}))$.

For example, if $n = 3$, then $v^1 = v(\tau^1, \tau^2, \tau^3)$, $v^2 = v(\tau^2, \tau^1, \tau^3)$, $v^3 = v(\tau^3, \tau^1, \tau^2)$, and $v(\tau^1, \tau^2, \tau^3) = v(\tau^1, \tau^3, \tau^2)$ etc. We can now use the function v (or, more precisely, functions based on it) to formulate the tariff reduction game precisely. As we are focussing on tariff reductions, we will assume throughout that $\boldsymbol{\tau} = (\tau^1..\tau^n) \in [0, \hat{\tau}]^n = F^n$.

From now on, for all $\tau, \tau' \in \mathcal{R}_+$, let $w(\tau, \tau') \equiv v(\tau, \tau', ...\tau')$ so $w(\tau, \tau')$ is any country $i's$ payoff in the event that i sets τ, and all $j \neq i$ set τ'. Without much loss of generality, we will assume that w is twice continuously differentiable i.e. w_1, w_2 be the partial derivatives of w with respect to τ, τ' respectively. 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 other countries' tariffs are below Nash equilibrium, any country likes an increase in its own (common) tariff, and a reduction in the tariffs of the other countries. 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 $\tau > 0$.

[9]This result, and all others, are proved in the Appendix, where a proof is required.

This says that any equal reduction in all tariffs, starting from a situation of equal tariffs 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, the closer other countries' tariffs are to Nash equilibrium tariffs, the smaller the gain any country makes from increasing its own tariff.

Payoffs over the infinite horizon are discounted by a common discount factor δ, $0 < \delta < 1$ i.e.

$$(1-\delta)\sum_{t=1} \delta^t w(\tau_t^i, \tau_t^{-i}) \tag{2.4}$$

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

Given the symmetry of the model, we restrict our attention to *symmetric* equilibrium[10] tariff paths where $\tau_t^i = \tau_t$, $t = 1, 2, \ldots$, i.e. where all countries choose the same tariff in every time period, and we denote such paths by the sequence $\{\tau_t\}_{t=1}^{\infty}$.

2.3. Withdrawal of Equivalent Concessions

Suppose that $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ is a candidate for an equilibrium tariff sequence, where $\tilde{\tau}_t$ is the tariff "agreed" for period t. Note that there are two kinds of punishment that $i \neq j$ could levy on j for deviating from $\{\tilde{\tau}_t\}_{t=1}^{\infty}$. One is to raise tariffs to the Nash level $\hat{\tau}$, the most severe credible punishment (which we call an *unconstrained* punishment). The other type of punishment is where $i \neq j$, upon observing that j has deviated at time $t-1$, withdraw precisely the equivalent concessions to market access at time t. That is, if the deviant j has set $\tau_{t-1}^j = \tau' > \tilde{\tau}_{t-1}$, then in the next period instead of retaliating by setting $\hat{\tau}$ the other parties withdraw the concessions made, implementing $\tau_t^i = \tau' = \tau_{t-1}^j$

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

as well. We call this form of punishment payoff a *withdrawal of equivalent concessions* (WEC). In practice, WTO members are bound by GATT/WTO rules to adopt exactly this penalty structure. To support WEC as a subgame perfect equilibrium punishment strategy, there *must* exist an implicit cost to a country of breaking the WEC (i.e. by setting some $\tau_t^i > \tau_{t-1}^j > \bar{\tau}_{t-1}$ in retaliation). Otherwise, it would never be observed to hold in practice. We denote this cost by c_i. Thus we have a stylized characterization of the WTO rules on withdrawal of equivalent concessions.[11] Finally, we assume that c_i is so high that no country would wish to violate the WEC rule. Given this, it is clear that the worst credible punishment that the set of countries $N/\{j\}$ can impose on j is to match the deviator's tariff in all subsequent periods.

3. Symmetric Equilibrium Paths

3.1. Optimal Deviations

We begin by characterizing the optimal deviation from a symmetric equilibrium path $\{\bar{\tau}_t\}_{t=1}^{\infty}$ for any country i, given that it rationally anticipates that it will be punished by the WEC rule. That is, all $j \neq i$ will match i's deviation tariff in all subsequent periods if and only if i deviates by setting a tariff $\tau' > \bar{\tau}_t$. Let i's optimal deviation at t from the reference path $\{\bar{\tau}_t\}_{t=1}^{\infty}$ be denoted z_t.

Note that the withdrawal of equivalent concessions applies only to deviation by setting a tariff *above* the agreed rate $\bar{\tau}_t$. Thus there is an asymmetry in the penalty. Formally, the payoff any country can expect from a deviation to z_t is:

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

We are interested in the *optimal deviation* z_t i.e. the choice of z_t that maximizes $\Delta(z_t, \{\bar{\tau}_t\}_{t=1}^{\infty})$ given the reference path. Due to the discontinuous nature of the payoff $\Delta(z_t, \{\bar{\tau}_t\}_{t=1}^{\infty})$, an optimal deviation does not exist, but we can precisely bound the gain from deviation.

[11] Elsewhere in the literature, reputation effects are modelled explicitly (e.g. Kreps, Milgrom, Roberts and Wilson 1982, Kreps and Wilson 1982, Milgrom and Roberts 1982). Here they are simply introduced by assumption as an enforcement device because we want to focus on the effect of the WEC penal code itself.

Technically, the largest possible gain from deviation is the supremum of $\Delta(z_t, \{\mathbf{e}_t\}_{t=1}^{\infty})$ across all values of $z_t \neq \mathbf{e}_t$, which we denote by $\overline{\Delta}(\{\mathbf{e}_t\}_{t=1}^{\infty})$.

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

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

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 above the agreed level; $z_t \geq \tilde{\tau}_t$. It can never benefit by a unilateral deviation $z_t < \tilde{\tau}_t$.[12] Now, from the first term in curly brackets which gives the gains to deviation, define

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

$z(.)$ can be thought of as a kind of "reaction function" indicating how the optimal deviation varies with the agreed tariff τ_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. We can think of $\zeta(\tau)$ as a kind of reaction function. 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, note that if A3 holds, $\zeta'(\tau) < 0$. 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 $z(.)$:

[12] 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 therfore ignore the possibility that $z_t < \tau_t$ because, by A1, a country would make itself worse off by deviating in this way.

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

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

3.2. Efficient Equilibrium Paths

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:

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

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 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 (2.4). 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. :

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

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(s) that are of most interest. We now turn to characterizations of the efficient equilibrium path. Our first main result, Proposition 2, establishes that free trade is in fact impossible under WEC.

Proposition 2. *(Failure to reach free trade) 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 all other countries agree to adopt free trade at *any* point in time, then the last will have an incentive to deviate by levying a positive tariff. So such an agreement would not be self-enforcing. 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 $\delta > \delta_0$. Instead, Proposition 3 is reminiscent of the results of Lockwood and Thomas(2002), who study a repeated prisoner's dilemma with *complete* irreversibility of actions.

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[13]. 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

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

[13]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 defintiion of α binds, implying that $z(\hat{\tau}) = \hat{\tau}$, and consequently, that $\alpha(\hat{\tau}) = (1-\delta)\psi(\hat{\tau},\hat{\tau}) + \delta\psi(\hat{\tau},\hat{\tau}) = \psi(\hat{\tau},\hat{\tau}) = \beta(\hat{\tau})$.

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 3. *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 3 shows that under a withdrawal of equivalent concessions it is possible for all 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.

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 4. *The stationary path, which has $e_t = \overline{\tau}$, 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).

We now illustrate our results with the quasi-linear example i.e. we assume that preferences take the form (2.2). This example is analyzed thoroughly in the appendix. First, it can be shown that the Nash equilibrium tariff is $\hat{\tau} = 1/(\sigma - 1)$. Also, we show 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, $\overline{\tau} \rightarrow \hat{\tau}$ as $\delta \rightarrow 0$, and $\overline{\tau} \rightarrow 0$ as $\delta \rightarrow 1$. 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 $\overline{\tau}$.

If the GATT/WTO provides a means by which countries select the efficient tariff reduction path, then Propositions 2, 3 and 4 provide a complete characterization of this path. Accordingly, under WEC trade liberalization can be achieved, but that free trade cannot be reached. However, at present our model cannot "explain" the gradualism in tariff-cutting observed in practice.

4. Loss of Political Good Will and Gradual Tariff Reduction

In Section 2.3, we argued that there must exist an implicit cost to countries of breaking the WEC penal code. If not, then it would never actually be observed to hold. This cost was posited as a loss of political good will, which would be exerted in other areas or the international political arena. This loss of political good will is now extended to the initial deviant. Specifically, we will assume the following. If country i sets $\tau_t > \tau_{t-1}$, it incurs a *political* cost of deviation $\mathbf{e}_i$.[14] If on the other hand $\tau_t \leq \tau_{t-1}$, country i incurs no such cost at the initial deviation.[15]

A justification of this penalty structure is as follows. Article 2 of GATT (1994) in the Charter of the WTO specifies that a schedule of commitments be maintained. Results of tariff negotiations are dutifully recorded as scheduled commitments in the form of tariff bindings; a permanent and irrevocable commitment that tariffs will *not rise above bound levels* for the product in question. Violations of tariff bindings become the subject of dispute settlement; with initial complaint, investigation and hearing before panels, panel findings, and rulings by the WTO council to come into compliance. Failure to return to compliance will eventually lead to retaliation being sanctioned by the WTO on the part of parties affected by the violation of bindings against violators.

Why has this been so? Why have tariff bindings under GATT/WTO de facto become

[14] We do not assume in general that $c_i = \mathbf{e}_i$. For example, it may be that the political cost of reneging on the original agreement in the first place is higher than the cost of deviating later, in the punishment phase. Or there may be a higher cost to losing the moral high ground.

[15] Note that a country can deviate from the agreement without incurring a loss of political good will by setting τ' so that $\mathbf{e}_t < \tau' < \tau_{t-1}$.

permanent and irreversible commitments, and what has been the penalty structure to maintain this system? Firstly, tariff bindings have acquired the status of an international commitment comparable to that of other international treaties. Bindings, if committed to, effectively slot into a box of enshrined cross country commitments comparable to military and diplomatic treaties. Violation of tariff bindings brings into question the soundness of a country's financial commitments, its trustworthiness in strategic and military matters, its diplomatic reputation. Violating tariff bindings has large costs outside the tariff area (Keohane 1982, 1984 chapter 4, Jackson 1989 chapters 2, 4).[16]

It is somewhat unsatisfactory that these political costs of tariff reversals are not firmly micro-founded. However, it appears that such costs exist and are important in the international arena. And no theory exists of which we are aware to explain the impact on tariff reductions of this type of cost. Therefore, in the absence of such a theory, it seems appropriate to simply assume that such costs exist in order to examine their consequences.

We will assume in what follows that $\mathbf{e}_i$ is high enough so that a deviation at t will never be above τ_{t-1} and thus incur loss of political goodwill. We can now reformulate the equilibrium conditions (3.5) under this new constraint. 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

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

As before, let the set of equilibrium tariff paths be E, and define the efficient tariff paths in E as those paths that maximize (2.4). Also as before, any efficient path must satisfy (4.1) with equality.

To proceed, we first introduce the following result. By Lemma 2, we know that $z(\tau_t) \geq \overline{\tau}$ for all $\tau_t < \overline{\tau}$.. So, $z(\tau_t) > \tau_{t-1}$ also if $\tau_{t-1} < \overline{\tau}$. Formally:

Lemma 3. *If* $\tau_t, \tau_{t-1} \leq \overline{\tau}$, *then* $\chi(\tau_t, \tau_{t-1}) = \tau_{t-1}$.

[16] Current (at the time of writing) protectionist measures imposed on steel imports by the European Union and US, appear to be in breach of tariff bindings. Yet over the postwar period in general, the main focus of this paper, instances of violations of tariff bindings are rare.

This says that as long as $\tau_t, \tau_{t-1} \leq \overline{\tau}$, the optimal retaliation is τ_{t-1}. Recall from Lemma 2 that $\zeta(\mathbf{e}) > \overline{\tau}$ if $\mathbf{e} < \overline{\tau}$. But now a loss of political good will prohibits a deviation to this level because $\zeta(\mathbf{e}) = z(\tau_t) > \overline{\tau} > \tau_{t-1}$. If the cost from a loss of political good will is high enough, the country is better off adopting τ_{t-1} rather than $\zeta(\mathbf{e}) = z(\tau_t)$; ie $\chi(\tau_t, \tau_{t-1}) = \min\{z(\tau_t), \tau_{t-1}\} = \tau_{t-1}$.

Now, suppose that $\{\tau_t\}_{t=s}^{\infty}$ is an efficient path from s onwards with $\tau_t \leq \overline{\tau}$, $t \geq s$. From (4.1) and Lemma 3, this path must satisfy

$$\begin{aligned}(1-\delta)(w(\tau_t,\tau_t) + \delta w(\tau_{t+1},\tau_{t+1}) + \ldots\ ) &= \\ (1-\delta)w(\tau_{t-1},\tau_t) + \delta w(\tau_{t-1},\tau_{t-1}),\ t &= 1, \ldots\end{aligned} \tag{4.2}$$

Advancing (4.2) one period, multiplying both sides by δ, subtracting from (4.2), and dividing the result by $1-\delta$, we get:

$$w(\tau_t,\tau_t) = w(\tau_{t-1},\tau_t) + \frac{\delta}{1-\delta} w(\tau_{t-1},\tau_{t-1}) - \delta \quad w(\tau_t,\tau_{t+1}) + \frac{\delta}{1-\delta} w(\tau_t,\tau_t) \tag{4.3}$$

which is a second-order difference equation[17] in τ_t. This can be seen more clearly by rearranging (4.3) to get:

$$w(\tau_t,\tau_{t+1}) = \frac{1}{\delta}\left[w(\tau_{t-1},\tau_t) - w(\tau_t,\tau_t)\right] + \frac{w(\tau_{t-1},\tau_{t-1})}{1-\delta} - \frac{\delta w(\tau_t,\tau_t)}{1-\delta},\ t > 1. \tag{4.4}$$

Let $\{\tau_t(\tau_0,\tau_1)\}_{t=2}^{\infty}$ be the sequence that solves (4.4) 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. *Any sequence $\{\tau_t(\tau_0,\tau_1)\}_{t=2}^{\infty}$ that satisfies (4.4), with initial conditions τ_0, τ_1 with $0 < \tau_1 < \tau_0$ is strictly decreasing i.e.* $0 < \tau_{t+1}(\tau_0,\tau_1) < \tau_t(\tau_0,\tau_1)$ all $t \geq 1$.

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) $\tau_1 \leq \overline{\tau}$ (required by Lemma 3: otherwise, the efficient path does not satisfy (4.4)). So, it remains

[17] 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 (2.4).

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.

$$\begin{aligned} & (1-\delta)(w(\tau_1, \tau_1) + \delta w(\tau_2(\hat{\tau}, \tau_1), \tau_2(\hat{\tau}, \tau_1)) + ...) \\ = \; & (1-\delta) w(\chi(\tau_1, \hat{\tau}), \tau_1) + \delta w(\chi(\tau_1, \hat{\tau}), \chi(\tau_1, \hat{\tau})) \end{aligned} \tag{4.5}$$

We now have:

Proposition 5. *There exists a smallest value of* τ_1, $0 < \tilde{\tau}_1 < \overline{\tau}$ *that satisfies (4.5). 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 5 we learn that 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 $t = 1, ..., s-1$ 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 future periods, then country i may do better by maintaining $\tilde{\tau}_s$ in $s+1$ whilst all other countries proceed to set $\tilde{\tau}_{s+1} < \tilde{\tau}_s$, even if all countries impose the WEC penal code in all periods after that. But it is always possible to promise additional reductions in future periods that can compensate for the gains to deviation in period s.

Why is the cost from loss of political good will necessary for this process? In its absence, the unilateral gains from deviating to $\overline{\tau}$ are greater than the gains from all future reductions. Indeed, the gains from deviation grow with the size of the overall reduction. But if a loss of political good will limits a deviation to the tariff level in the previous period, $\tilde{\tau}_{t-1}$, then the promise of all future reductions *can* be large enough to compensate for the gains from deviation in a single period.

Proposition 5 establishes that we can restrict attention to a tariff reduction sequence $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ for which $0 < \tilde{\tau}_t \leq \overline{\tau}$ all $t \geq 1$. We are able to explore the properties of the efficient equilibrium tariff reduction path further by looking at specific examples. To do this, the functional form must be specified and it must be verified that for the example under consideration assumptions A1-A3 are satisfied.

4.1. Computed Paths

A computational algorithm is used to find the efficient equilibrium path. This entails finding the smallest possible value of τ_1 - $\tilde{\tau}_1$ - that satisfies (4.5) and therefore, by Proposition 5, gives rise to the unique efficient path $\tilde{\tau}_t = \tau_t(\hat{\tau}, \tilde{\tau}_1)$. There exists no analytical way of finding $\tilde{\tau}_1$, but it can be approximated in the following way. First, set the second initial condition, τ_1, of the difference equation defined in (4.5) equal to the efficient stationary tariff, $\overline{\tau}$. (Recall that the first condition is fixed at $\tau_0 = \hat{\tau}$.) Then reduce this second initial condition by a small step ε and check that the resulting difference equation converges to some positive tariff rate. Continue in this way, reducing τ_1 by steps of ε until it is so low that the difference equation diverges. The final convergent difference equation is then the approximation to the efficient path. The approximation is more accurate the smaller the step size ε. Intuitively, the efficient tariff reduction path cannot bring about non-positive tariffs, because free trade cannot be reached, by Proposition 2.

The algorithm is as follows:

1. Let $k = k + 1$.

2. Set $\tau_0 = \hat{\tau}$ and $0 < \tau_1 = \overline{\tau} - k\varepsilon \leq \overline{\tau}$ as initial conditions and solve (4.4) forward for T periods.

3. If $\tau_T(\tau_1; \tau_0, \delta) > 0$, set $S_k = S_{k-1} \cup \{\hat{\tau} - k\varepsilon\}$ and go to 1.

4. If $\tau_T(\tau_1; \tau_0, \delta) \leq 0$, stop. Discard this path.

This algorithm is initialized by setting $S_0 = \varnothing$. Note that the algorithm can only run at most for m steps, where m is the largest integer smaller than $\overline{\tau}/\varepsilon$. Let $K + 1 \leq m$ be the number of steps after which the algorithm stops. The algorithm stops when a path fails the criterion of $\tau_T(\tau_1; \tau_0, \delta) > 0$. Having failed, this last path must be discarded. Then $S_K = S_{K-1} \cup \tau_1 - K\varepsilon$ and $\tau_1 = \overline{\tau} - K\varepsilon$ is the smallest member. S_K then comprises the full set of tariff reduction paths that satisfy (4.5), and $\tau_1 = \overline{\tau} - K\varepsilon = \tilde{\tau}_1$ gives rise to the efficient path, as required.

The technical details are as follows. The utility function (2.2) is substituted into the second order difference equation that defines an equilibrium tariff reduction path (4.4).

The resulting expression is used to solve sequentially for the equilibrium tariff level τ_{t+1}, given levels in τ_{t-1} and τ_t. Recall that the algorithm requires the size of the steps between simulations ε and the total number of periods T to be determined. We use, respectively, $\varepsilon = 0.0001$ and $T = 10000$. A smaller value of ε and a larger value of T would yield greater accuracy in computation of the equilibrium reduction path, but take longer.

The procedure is begun with $k = 0$, so in calculating S_0 the procedure is initialized using $\tau_0 = \hat{\tau}$, $\tau_1 = \overline{\tau}$. Let K be the highest value of k for which $\tau_T(\tau_1; \tau_0, \delta) > 0$. The algorithm is illustrated in Figure 3, for $\sigma = 2$ and $\delta = 0.5$, where the path corresponding to step $k = K$ is the approximation to the efficient tariff reduction path. The tariff level is shown on the vertical axis, with simulation periods on the horizontal axis. Only the first 1000 periods of the simulation are presented. We also show what happens for $k = K + 1$ and $k = K + 2$. Note that no value for the number of countries is specified. The reason is that n has no impact whatever on the equilibrium path under the quasi-linear preference specification.[18]

Given $\sigma = 2$, $\delta = 0.5$, and $\tau_0 = \hat{\tau} = 1$ we have $\tau_1 = \overline{\tau} = 0.25$ for $k = 0$ and $\tau_1 = 0.2499$ for $k = 1$ and so on. One of the paths shown in Figure 3 is for $k = K = 1426$, so that $\tau_1 = 0.1704$. Note that for this set of initial conditions, the reduction path stabilizes; $\tau_{10000} = 0.102748 > 0$. This is the efficient gradual reduction path. How do we know? When k is increased by 1 to $K + 1 = 1427$, the criterion $\tau_T(\tau_2; \tau_1, \delta) > 0$ fails.

This path that fails the criterion is also presented in Figure 3. Observe that $k = K + 1 = 1427$ implies $\tau_1 = 0.1703$. The path diverges sharply downwards and τ_{10000} - were it to be displayed - would be significantly below 0, failing the criterion for that path to be an equilibrium. At $t = 100$, $\{\tau_{100}(\overline{\tau} - (K+1)\,\varepsilon; 1, 0.5)\} = 0.099384$, and is close to $\{\tau_{100}(\overline{\tau} - K\varepsilon; 1, 0.5)\}$. However, as t increases further the path of the sequence $\{\tau_t(\overline{\tau} - (K+1)\,\varepsilon; 1, 0.5)\}_{t=1}^{T}$ diverges downwards sharply from $\{\tau_t(\overline{\tau} - K\varepsilon; 1, 0.5)\}_{t=1}^{T}$, so $\tau_T(\tau_1; \tau_0, \delta) \leq 0$ for $K + 1$ and the path must be discarded (see Step 4 of the algorithm above). For $K + 2$, where $\tau_1 = 0.1702$, the divergence takes place at an even lower value of t.

Figure 3 also shows the one off tariff reduction path, with the tariff being reduced

[18] To put this another way, if a closed form solution for the reduction path could be found, then n would cancel from the expression.

immediately to $\overline{\tau}$ in period 1. Between this tariff and the most efficient tariff reduction path lies the 'Region of gradual reduction paths' which (in the limit) fills the area between the one off reduction path and the efficient gradual reduction path.

On a cautionary note, the algorithm may pick a path that appears to approximate the equilibrium path for a given value of T, but fails for some larger T. In view of this possibility the value of K and corresponding τ_1 for the optimal path given here by $\tau_1 = 0.1704$ was checked for robustness by setting $T = 100000$ and verifying that $\tau_T(\tau_1; \tau_0, \delta) > 0$ continued to hold. The same robustness check was also performed on all other computed optimal paths presented below.

Figures 4 and 5 illustrate efficient tariff reduction paths that result from comparative dynamics exercises carried out using the quasi-linear preference function on the same format as Figure 3. These latter figures present only the first 250 of 10000 periods. Figure 4 shows how the optimal reduction path varies with the substitution elasticity σ, whilst Figure 5 indicates the impact of variation in the discount factor δ.

Look at Figure 4 first. There are optimal reduction paths for three substitution elasticities $\sigma = 2$, 5 and 10 with the other parameter held fixed at $\delta = 0.5$. The key data and results for these simulations are presented in boxes on the far right hand side of the figure. As in Figure 3, for each value of σ we already know $\hat{\tau}$ and $\overline{\tau}$ from the analysis. Both are decreasing in σ, and the figure shows that the optimal reduction paths are monotonically decreasing in σ as well, as one would expect.

The discount rate δ only affects the reduction path, and not $\hat{\tau}$, explaining why the optimal reduction paths in Figure 5 start at the same point and decline towards different limits. Simulations for $\delta = 0.1$, 0.5 and 0.9 are shown, holding $\sigma = 2$ constant. We see that for higher values of δ the liberalization path exhibits greater liberalization at each point in time t.

5. Conclusions

This present 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 the institutional structure set up in the framework of the GATT to achieve trade liberalization, paying special attention to the role of time in the process. We use a dynamic game framework, which makes it is possible to take account of the fact that a country is able to renege on an agreement for some time before being found out. In addition, the GATT/WTO institutional structure limits the extent of allowable retaliation. It is the interaction of these two features in our model, novel in the present context, which enables us to explain the failure to reach free trade and gradualism.

We return to an apparent difference in the outcome from our modelling framework to that of Bagwell and Staiger (1999). They also model a trade agreement using a penalty structure based on the WTO's withdrawal of equivalent concessions as a penalty structure. However, in their model it is possible to achieve full efficiency whilst in ours it is not. In their conclusion, they point out that there may in fact be enforcement difficulties. (As Bagwell and Staiger point out, enforcement difficulties have been studied in a wider context by Dam 1970). Our dynamic game captures and formalizes an element of this enforcement difficulty that Bagwell and Staiger' model does not; that a country is able to reap the benefits of deviation for a period before retaliation occurs. It is this that drives the inability to obtain full efficiency in our model, which is not a feature of Bagwell and Staiger's.

Inevitably, the theoretical framework simplifies the situation in a number of key respects. All countries are assumed to be symmetrical, and small in terms of their purchasing power on world markets relative to the political costs of raising protectionism. Each country exports only a single good, with all 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 perpetual liberalization considerably more subtle and complex, with more rapid progress achieved in areas where countries receive greater gains from protectionism relative to the political costs incurred. 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 WTO penalty structure that we examine here. By defining a symmetrical modelling framework this issue is completely suppressed in our present paper.

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. It may be that our framework provides 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 be protectionist in order to gain financial backing from industrialists (Grossman and Helpman 1995) and for electorates to elect politicians who signal that they will adopt protectionist measures in order to increase their chances of being elected (Riezman 2001). These protectionist forces may be outweighed at an early stage by the gains that we describe which are relatively large early on in the process, but not later once the potential gains become relatively small. Future research could study the interaction of these counteracting forces.

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 (1999); "An Economic Theory of the GATT." *American Economic Review*, 89(1) (March): 215-48.

[3] Bond, E.B. and J. H. Park (2000); "Gradualism in Trade Agreements with Asymmetric Countries." To appear in *Review of Economic Studies.*

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

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

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

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

[8] 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.

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

[10] General Agreement on Tariffs and Trade (GATT 1994); "Uruguay Round: Final Act" Marrakesh, 15 April. Photocopy.

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

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

[13] Jackson, J. (1989); *The World Trading System.* MIT Press, Cambridge Massachusetts, 417 pages.

[14] Keohane, R. (1982); "The Demand for International Regimes." *International Organization* 36(2) (Spring) 325-355.

[15] Keohane, R. (1984); *After Hegemony*: *Cooperation and Discord in the World Political Economy.* Princeton University Press, Princeton, New Jersey.

[16] Kreps, D., P. Milgrom, D. Roberts and R. Wilson (1982); "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." *Journal of Economic Theory*, 27, 253-279.

[17] Kreps, D. and R. Wilson (1982); "Reputation and Imperfect Information." *Journal of Economic Theory*, 27, 253-279.

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

[19] Lockwood, B. and J. Thomas (2002); "Gradualism and irreversibility." To appear in *Review of Economic Studies.*

[20] 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.

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

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

[23] Milgrom, P. and D. Roberts (1982); "Predation, Reputation and Entry Deterrence." *Journal of Economic Theory*, 27, 280-312.

[24] Riezman, R. (2001) "The sources of protectionist drift in representative democracies." typescript.

[25] 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.

[26] Syropoulos, C. (2001); "Optimum Tariffs and Retaliation Revisited: How Country Size Matters." To appear in *Review of Economic Studies.*

A. Appendix

A.1. Proof of Propositions

Proof of Proposition 1. Fix $i \in N$, and normalize prices by setting $p_i = 1$, so $\mathsf{p} = (p_1, ..p_{i-1}, p_{i+1}, ..p_n)$. Then, by the symmetry of the model, and taking τ_i as fixed,

$$\mathsf{p}(\tau^i, \pi(\tau^{-i})) = \pi(\mathsf{p}(\tau^i, \tau^{-i})),\ R^i(\tau^i, \tau^{-i}) = R^i(\tau^i, \pi(\tau^{-i})) \quad \text{(A.1)}$$

where $\pi(.)$ is any permutation function i.e. a permutation in tariffs of other countries leads to the same permutation in their equilibrium prices, as tariffs are the only variables affecting excess demands that differ across countries. Now note that by definition,

$$v^i(\tau^i, \tau^{-i}) \equiv v^i(\mathsf{p}(\tau^i, \tau^{-i}), \tau^i, R^i(\tau^i, \tau^{-i})) \quad \text{(A.2)}$$

Also, by symmetry of the model,

$$v^i(\pi(\mathsf{p}(\tau^i, \tau^{-i})), \tau^i,) = v^i(\mathsf{p}(\tau^i, \tau^{-i}), \tau^i, R^i) \quad \text{(A.3)}$$

i.e. country utility is the same if the world prices of imports are permuted. So we have

$$\begin{aligned} v^i(\tau^i, \pi(\tau^{-i})) &= v^i(\mathsf{p}(\tau^i, \pi(\tau^{-i})), \tau^i,\ R^i(\tau^i, \pi(\tau^{-i}))) \quad \text{(A.4)} \\ &= v^i(\pi(\mathsf{p}(\tau^i, \tau^{-i})), \tau^i,\ R^i(\tau^i, \tau^{-i})) \\ &= v^i(\mathsf{p}(\tau^i, \tau^{-i}), \tau^i,\ R^i(\tau^i, \tau^{-i})) \\ &= v^i(\tau^i, \tau^{-i}) \end{aligned}$$

where the first line of (A.4) is from (A.2), the second is from (A.1), the third is from (A.3), and the fourth is from (A.2) again. This proves the second part of the Lemma. To prove the first part, note that as all countries are identical up to a permutation of the indices of the goods, $v^j = v^i(\tau^j, \tau^{-j})$, all i, j so $v^i = v(\tau^i, \pi(\tau^{-i}))$ as required. $\square$

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

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

(b) If a country deviates to $z_t > \tau_t$, it receives

$$g(z_t, \tau_t) = (1 - \delta)w(z_t, \tau_t) + \delta w(z_t, z_t) \tag{A.5}$$

So, it suffices to show that (A.5) 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 consumption bundle $\mathsf{x}(z^n, \tau_t)$ 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 $\mathsf{x}(\tau_t, \tau_t)$ generated by not deviating. Contradiction. ¤

Proof of Lemma 2. 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 $\zeta(\tau) > \tau$, $\tau < \overline{\tau}$, $\zeta(\tau) < \tau$, $\tau > \overline{\tau}$.

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 $\overline{\tau} = \zeta(\tau)$, we see that $z(\tau) = \zeta(\tau)$, $\tau < \overline{\tau}$, and $z(\tau) = \tau$, $\tau \geq \overline{\tau}$. ¤

Proof of Proposition 2. Suppose to the contrary that $\tau_t = 0$ for some t. Then, at t, the incentive constraint is

$$(1-\delta)w(0,0) + \delta w(0,0) \geq (1-\delta)w(z(0), 0) + \delta w(z(0), z(0)) \tag{A.6}$$

Now, we will show that at the solution to problem (3.2), $z(0) > 0$. It will then follow that

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

contradicting (3.5). 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 3. 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 $\tau^* < \overline{\tau}$.[19] Now, the slope of β is

$$\beta'(\tau) = w_1(\tau,\tau) + w_2(\tau,\tau) \tag{A.7}$$

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:

$$\alpha'(\tau) = (1-\delta)w_2(z,\tau) \tag{A.8}$$

Given $z \geq \tau$ in (A.8), 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

$$\alpha'(\tau) \leq (1-\delta)w_2(\tau,\tau). \tag{A.9}$$

So, from (A.7), (A.9), the required condition is that

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

[19] 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 $[\tau^*, \overline{\tau}]$.

Rearranging, this is

$$0 < w_1(\tau, \tau) + \delta w_2(\tau, \tau) \tag{A.10}$$

But, the FOC defining τ is:

$$w_1(\overline{\tau}, \overline{\tau}) + \delta w_2(\overline{\tau}, \overline{\tau}) = 0 \tag{A.11}$$

As $\tau < \overline{\tau}$, from (A.11) we must have:

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

where the derivatives on the RHS of (A.12) are evaluated at (x, x). By A3, $w_{12} < 0$. By assumption, $w_{11}, w_{22} \leq 0$. So, (A.12) implies (A.10), as required.

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

Proof of Proposition 4. (a) Following the proof of Lockwood and Thomas (2002), Lemma 2.2, the equilibrium conditions (3.6) can be shown to be equivalent to the following difference equation,

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

with initial condition $\tau_0 = \hat{\tau}$, plus the condition that the solution to (A.13) is bounded. To see this, note first that advancing the equality in (A.13) 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:

$$\begin{aligned}(1-\delta)w(\tau_t, \tau_t) &= (1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t)) \\ -\delta&\left[(1-\delta)w(z(\tau_{t+1}), \tau_{t+1}) + \delta w(z(\tau_{t+1}), (z(\tau_{t+1}))\right],\ t = 1, ..\end{aligned} \tag{A.14}$$

Using the definitions of α, β in (A.14) and rearranging, we get[20] (A.13).

(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).

[20] The converse result can be obtained by solving (A.13) 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.13) implies (3.6).

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

$$\alpha(\tau_{t+1}) \leq \alpha(\tau_t) \tag{A.15}$$

Combining (A.13) and (A.15), 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.13) 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. ¤

Proof of Lemma 4. The proof is by induction. Assume $\tau_t < \tau_{t-1}$. Rewriting (4.4), we get:

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

By Lemma 3,

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

where the third line follows by definition. And because $0 < \delta < 1$, it follows that $\delta\left[w(\tau_t,\tau_{t+1}) - w(\tau_t,\tau_t)\right] > 0$. So, $w(\tau_t,\tau_{t+1}) > w(\tau_t,\tau_t)$. But then, by A1, $\tau_{t+1} < \tau_t$, as required. ¤

Proof of Proposition 5. First, rewrite (4.5) as a function of τ_1 :

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

Now, note that by the definition of $\overline{\tau}$,

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

Moreover, $\tau_t(\hat{\tau},\overline{\tau})<\overline{\tau}$, all t by Lemma 4. So, if $\tau_1=\overline{\tau}$, (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 assume that the utility function is of quasi-linear form given by (2.2). Maximization of (2.2) subject to (2.3) gives demands for the two goods;

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

$$x_i^i = 1+\frac{R_i}{p_i}-\sum_{j\neq i}\frac{p_j(1+\tau_j^i)x_j^i}{p_i}=1+\frac{R_i}{p_i}-\sum_{j\neq i}\left(\frac{p_j(1+\tau_j^i)}{p_i}\right)^{1-\sigma} \tag{A.17}$$

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 (2.2) to get

$$v^i = \frac{1}{\sigma-1}\sum_{j\neq i}\left(\frac{p_j(1+\tau_j^i)}{p_i}\right)^{1-\sigma} + \frac{R_i}{p_i} \tag{A.18}$$

Also, tariff revenue is

$$R_i = \sum_{j\neq i} p_j\tau_j^i x_j^i = \sum_{j\neq i}\frac{p_j\tau_j^i}{p_i}\left(\frac{p_j(1+\tau_j^i)}{p_i}\right)^{-\sigma} \tag{A.19}$$

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

$$v^i = \frac{1}{\sigma-1}\sum_{j\neq i}\left(\frac{p_j(1+\tau_j^i)}{p_i}\right)^{1-\sigma} + \sum_{j\neq i}\frac{p_j\tau_j^i}{p_i}\left(\frac{p_j(1+\tau_j^i)}{p_i}\right)^{-\sigma} \tag{A.20}$$

Now, in Nash tariff equilibrium, a given country will always set the same tariff on all imported goods. So, we may suppose that all countries $j \neq i$ set a tariff $\tau' = \tau_k^j$ on imports from all countries $k \neq j$, and country i sets tariff $\tau = \tau_k^i,\ k \neq i$. Then, we only need to find the best response τ to τ' to characterize the Nash equilibrium in tariffs. If $\tau' = \tau_{jk},\ k \neq j, ..n,\ \tau = \tau_k^i,\ k \neq i$, then in equilibrium, $p_j = p$, all $j \neq i$. So, we may choose p_i as the numeraire. Using these simplifications, we may rewrite (A.20) as

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

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 $\tau' = \tau_{jk},\ k \neq j, ..n,\ \tau = \tau_k^i,\ k \neq i$, $p_j = p,\ j \neq i, p_i = 1$, we get;

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

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

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 ($1 = \sum_{i\in N} x_i^j$), we have

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

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

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

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,

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

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

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

Eliminating common terms, we get

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

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:

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

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

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

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.

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

Now A2:

$$\begin{aligned} &w_1(\tau, \tau') + w_2(\tau, \tau') = \\ &\qquad -(n-1)\,\frac{\sigma\,(1+\tau)^{-2-\sigma}\,(\sigma\tau\,(2+\tau+\tau') - (1+\tau)\,\tau')}{2\sigma-1}\left(\frac{1+\tau}{1+\tau'}\right)^{\left(1-\sigma-\sigma^2\right)/(1-2\sigma)} \end{aligned}$$

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:

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

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 $\overline{\tau}$. Dropping time subscripts and 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 (2.2) as follows

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

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

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

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[21] is

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

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

Figure 1

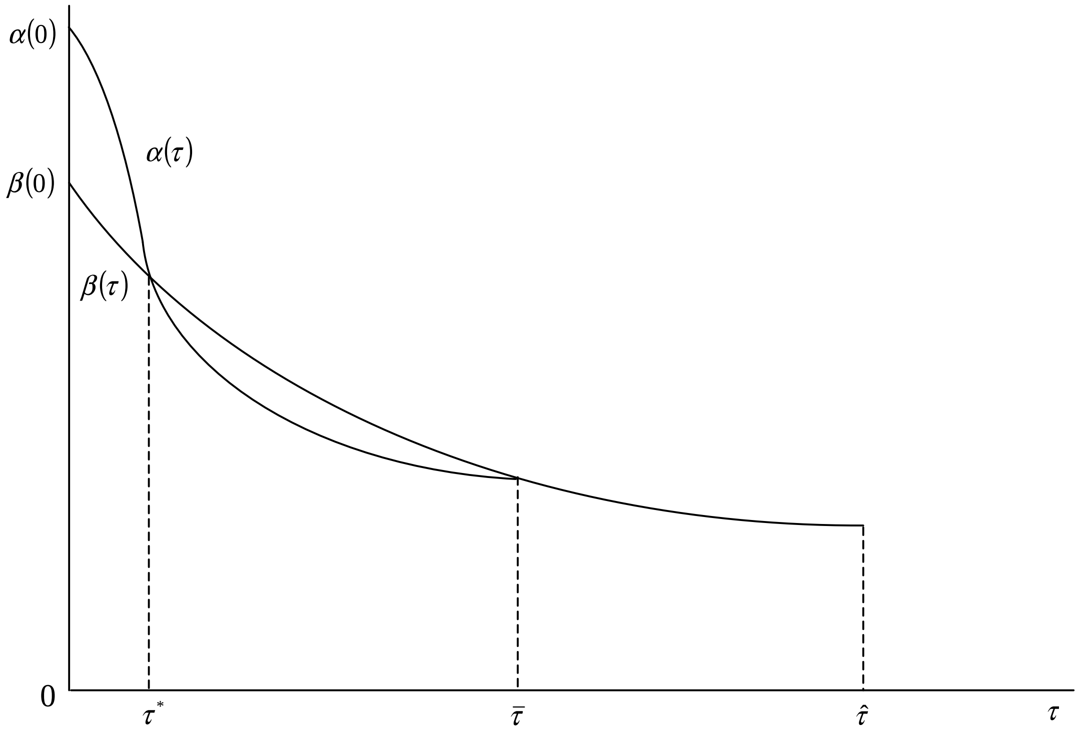

Figure 2

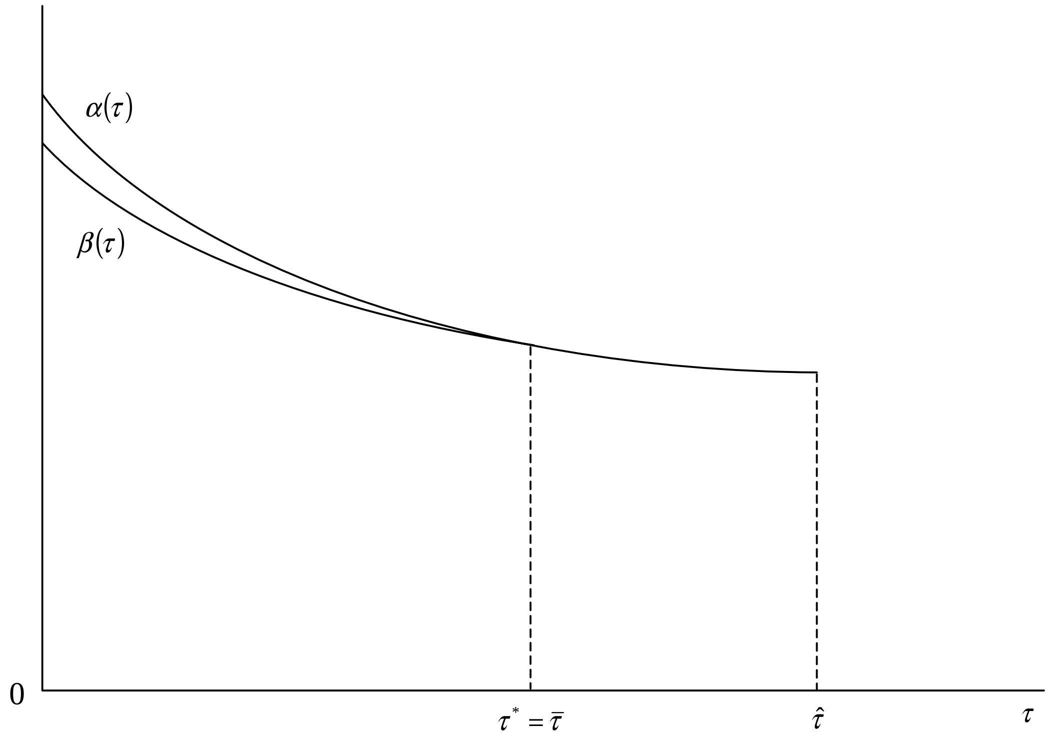

Figure 3; Approximating the Optimal Tariff Reduction Path

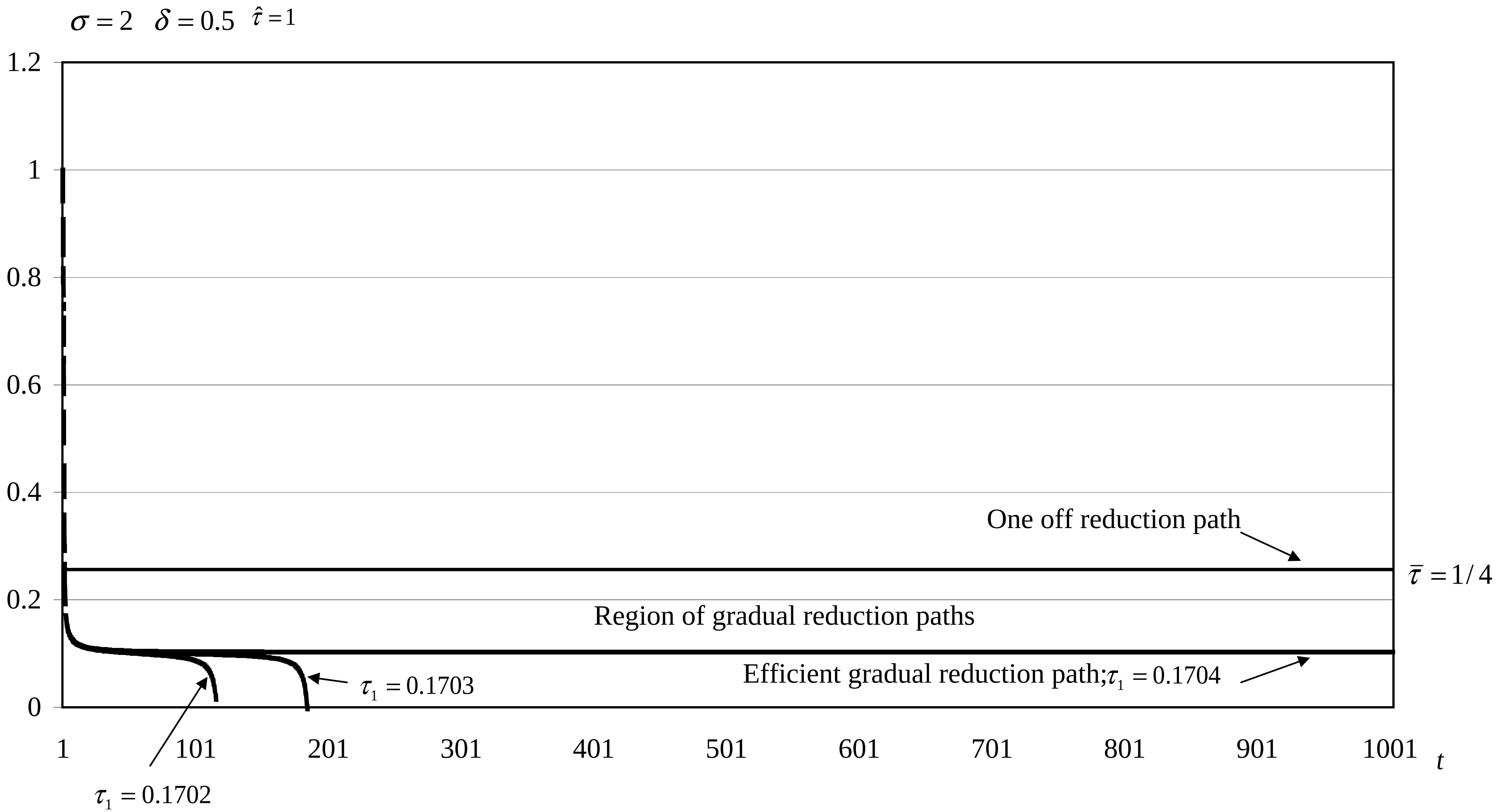

Figure 4: The approximate optimal tariff reduction path for various substitution elasticities

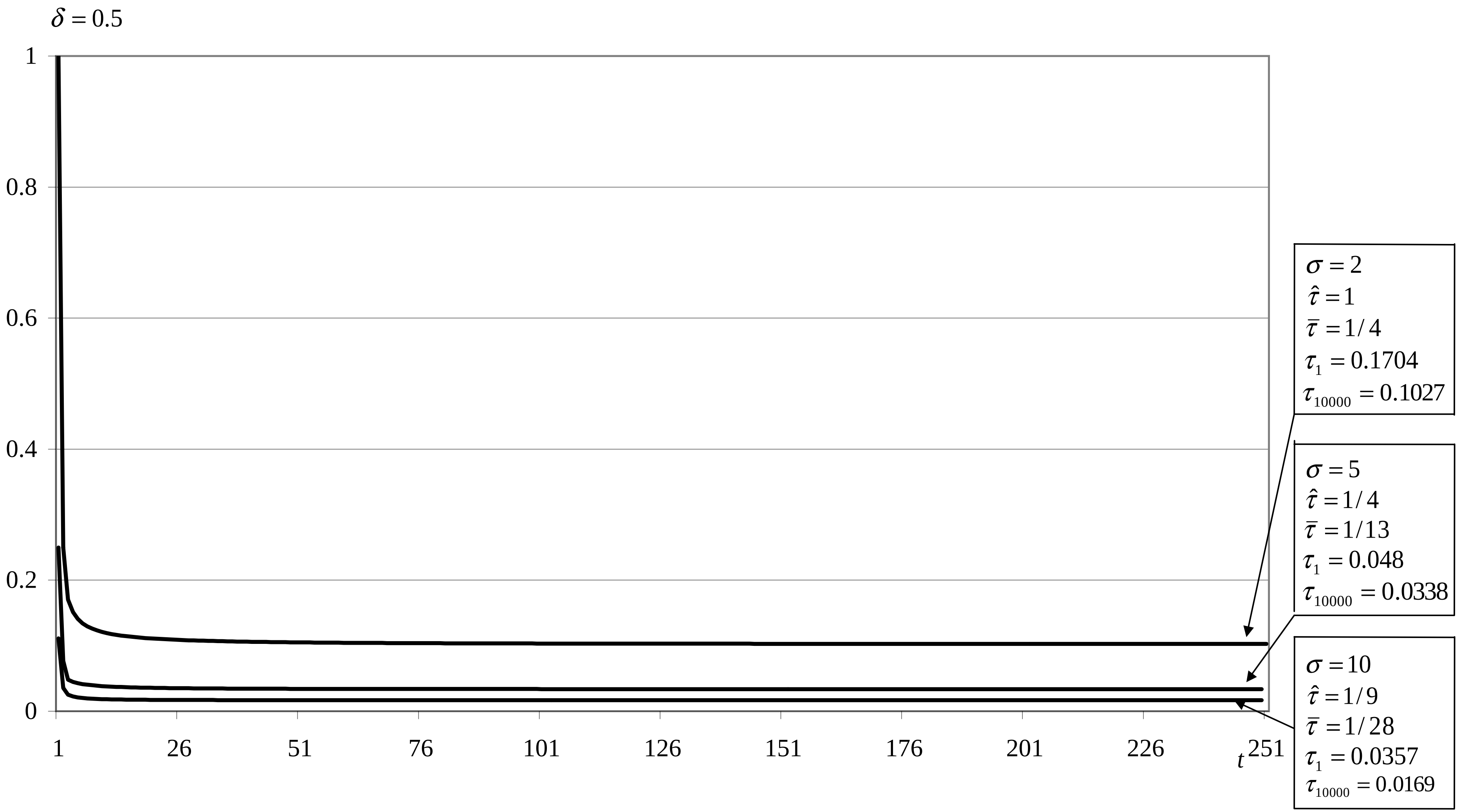

Figure 5: The approximate optimal tariff reduction path for various discount rates