ESSAYS ON UNCERTAINTY AND ESCAPE IN TRADE AGREEMENTS

by

Mathias Herzing

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

Monograph Series
No. 50

2005

THE INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
University of Stockholm

is an institute for independent scholarly research in the field of international economic relations.

MAIL ADDRESS: S-106 91 STOCKHOLM, SWEDEN
STREET ADDRESS: Universitetsvägen 10 A, 8th floor
TELEPHONE: + 46 8 16 20 00, **TELEFAX:** + 46 8 16 14 43

ESSAYS ON UNCERTAINTY AND ESCAPE IN TRADE AGREEMENTS

by

Mathias Herzing

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

February 2005

Institute for International Economic Studies
Stockholm University

ISBN 91-7155-000-3

Printed by Akademitryck AB
Edsbruk, Sweden 2005

Doctoral Dissertation
Department of Economics
Stockholm University

ABSTRACT

This thesis comprises three theoretical essays on uncertainty and escape in trade agreements.

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

Escape and Optimal Compensation in Trade Agreements addresses the issue of safeguard provisions in trade agreements which allow signatory countries to escape agreed-upon liberalization commitments under certain contingencies. In an infinitely repeated Prisoner's Dilemma tariff setting game between two countries, shocks influencing the incentive to deviate are introduced. Under asymmetric information about these shocks, liberalization is associated with a positive probability of cooperation breaking down in finite time. By introducing an escape clause allowing for temporary deviation while compensating the trading partner, cooperation can be sustained for any degree of trade liberalization. The compensation cost is shown to have an efficiency-enhancing effect by restraining the use of the escape clause. In fact, the expected per-period payoff increases for any given degree of liberalization if

the optimal fixed compensation cost scheme is implemented, as compared to the case when no escape clause is applied. Moreover, the scope for liberalization unambiguously increases in the presence of an escape clause under the optimal compensation cost scheme.

Optimal Time Limits on Safeguards in Trade Agreements addresses the issue of having time limits on how long countries should be permitted to withdraw liberalization commitments under a trade agreement. In a setting with two countries and an infinite number of sectors, each sector is subject to stochastic switches between two states over time. Under trade liberalization, there are gains to be made in the good state, while losses will be incurred when being in the bad state and protection by means of a safeguard is thus desirable. It is shown that, by limiting the time the safeguard can be applied, the interests of winners and losers in liberalization are balanced across countries. However, an ex ante agreed-upon finite time limit on the use of the safeguard will eventually be perceived as too short, as the share of sectors in need of being exempted from it increases over time. In the case when there is asymmetry between countries, a similar solution for the optimal time limit is obtained through Nash bargaining.

To My Parents

ACKNOWLEDGMENTS

My journey towards the PhD started some five years ago at the Department of Economics, and it now comes to an end at the Institute for International Economic Studies. It may be tempting to offer an accurate account of this journey with its ups and downs, but here, the focus will be on its brighter aspects. A number of people that have accompanied and truly helped me on this journey are the subject of my deep appreciation. First and foremost, I want to thank my excellent supervisors Harry Flam and Henrik Horn. I was extremely fortunate to encounter Henrik during my second year in the PhD program. Besides his personal qualities, his research interests appealed to me; I have benefitted enormously from having focused on topics within Henrik's area of expertise. For financial support during the last three years and a half, I thank Tom Hedelius' and Jan Wallander's Research Foundation. I also wish to thank Jaime Behar who through his supervision of my Master's thesis in the fall of 1998 made me return to academia, and moreover gave me the opportunity to be at the Latin American Institute at the beginning of the PhD program. Thanks also to the Department of Economics, where I spent the first two years of the program, in particular to Alberto Naranjo who started this journey with me and has been a friend ever since. I am indebted to the Institute for International Economic Studies and its staff, in particular to Christina Lönnblad and Annika Andreasson for editorial assistance and other administrative support. My appreciation also goes to my fellow graduate students, in particular Thomas Eisensee, Giovanni Favara, Natalie Pienaar and Ulrika Stavlöt, who all contributed to making academic life easier and more enjoyable. I am very grateful to my parents for their continuous support and encouragement. Most importantly, however, this journey would have been impossible without my beloved family, my wife Jai and our kids Wanna and Niclas.

TABLE OF CONTENTS

Chapter 1

Introduction

Trade agreements are typically meant to reduce trade barriers over an indefinite period of time. Their long-term nature can partly be explained by the necessity to give producers sufficient time to adjust their production to take advantage of the lower barriers, partly by negotiators' desire to create a stable political environment and partly by the very substantial costs of negotiating the highly complex agreements. There are not only advantages with long-term agreements, however. A potential drawback is that since the underlying economy is constantly changing, the agreement needs to be able to adapt. This would not be a problem in principle, if a sufficiently detailed state-contingent agreement could be struck. However, in practice, this is obviously not feasible; instead, agreements will have to be more or less rigid. But a completely inflexible agreement would not be very attractive. Countries would not be willing to commit to substantial irrevocable trade liberalization and, as a consequence, since trade agreements are not enforced by third parties but must be self-enforcing, such a completely rigid agreement would not be credible. Trade agreements therefore invariably contain mechanisms allowing for ex post flexibility. This thesis studies one such form of arrangement: so-called safeguards.

A safeguard is a provision permitting a signatory country to withdraw or cease to apply its normal obligations in order to protect certain overriding interests under specified conditions. Here, the focus will be on what is usually referred to as the "escape clause", Article XIX of the General Agreement on Tariffs and Trade, which is aimed at situations where a country suffers from sudden import surges that seriously threaten domestic industries. To prevent or remedy injury due to liberalization commitments, Article XIX provides the possibility of temporarily suspending obligations under the agreement.

Safeguard instruments can be seen to have two major functions. On the one hand, safeguards serve to alleviate unforeseen consequences of liberalization. On the other hand, by providing governments with the possibility for readjustments ex post, safeguards make it easier for countries to commit to liberalization ex ante. Hence, while the use of safeguards may threaten liberalization ex post, their presence makes it easier for countries to make liberalization commitments. In fact, the inclusion of safeguards in a trade agreement may lead to more far-reaching liberalization. The implications of safeguards for liberalization, both ex ante and ex post, are therefore important to study, in order to gain insight into how trade agreements should be designed.

This thesis comprises three essays. In very broad terms, these address (i) the effects of asymmetric information on the ability to maintain a self-enforcing trade agreement in a situation where the agreement cannot be state contingent; (ii) the benefits of including a safeguard provision in the presence of this type of uncertainty; and (iii) the role of limitations on the length of time safeguards are permitted to be applied.

The essay "Can Self-Destructive Trade Agreements Be Optimal?" examines the impact of uncertainty in the form of hidden information on the possibilities to sustain cooperation under a trade agreement. Two symmetric countries interact strategically through the setting of import tariffs in an infinitely repeated setting. This interaction is of a Prisoner's-Dilemma-type where it is a dominant strategy to deviate from any agreement to set tariffs below their Nash equilibrium levels, because deviation will yield a higher current-period payoff. But when the incentives to deviate facing the trading partner are perfectly observable, cooperation can nevertheless be sustained for sufficiently low degrees of liberalization, through the threat of punishing deviation in future periods.

It is, however, often the case that the incentives to deviate, weighing short-term gains against long-term losses, are not entirely known by other parties. A special case, where the weights governments attribute to present vis-à-vis future payoffs are stochastically determined and only privately observable, will be studied. In this case, both governments will have to infer the likelihood of cooperation being chosen by the trading partner. This inferred probability feeds back into the decision-making process such that a threshold value is obtained, prescribing under which realizations of the random variable adherence to the agreement will be chosen. These threshold values, one for each government, must be consistent in the sense that each country's threshold value is a best-response to the trading partner applying its threshold value. The consistency requirement will yield at least one threshold value solution for the weight attributed to present payoffs, such that cooperation will be chosen if and only if weight realizations are below that threshold value.

As customary in these types of trade agreement models, when at least some weight will always be attributed to future payoffs, cooperation can be infinitely sustained through the threat of infinite reversion to the suboptimal Nash equilibrium, if cooperative tariffs are sufficiently close to the Nash tariff level. If tariffs are further reduced, however, either cooperation breaks down instantly, or it can only be sustained as long as governments are sufficiently patient, with the likelihood of breakdown increasing as the cooperative tariff decreases. In the latter case, governments will thus ex ante have to trade off the benefit of farther-reaching liberalization against the cost of an increasing risk of cooperation breaking down. Or, in other words, governments will have to choose between safe agreements with low degrees of liberalization that can be sustained over the infinite horizon and self-destructive agreements with higher degrees of liberalization associated with strictly positive probabilities of cooperation breaking down in every period.

An interesting question is thus whether it may in fact be optimal to design the agreement such that it will break down in finite time? It is shown that whether liberalization can be reduced beyond the point where cooperation is always sustainable, and whether such high degrees of liberalization are optimal, solely depends on the properties of the density function of the stochastic variable. A self-destructive agreement will be the preferred option if countries are, on average, sufficiently impatient, or if the most myopic realization of the current-period payoff weight is sufficiently large, or if the marginal likelihood of its occurring is sufficiently low.

The above essay portrays the optimal trade agreement in a situation where a safeguard agreement is not permitted. But the possibility that the agreement may ultimately break down due to a lack of flexibility to adapt to exogenous shocks suggests that some form of safeguard may indeed have a role to play. This issue is addressed in the essay "Escape and Optimal Compensation in Trade Agreements". A similar setting to that in the first essay is employed here to examine the impact of a safeguard. In this infinitely repeated Prisoner's Dilemma tariff setting game between two symmetric countries, the one-period gain from deviating is again stochastically determined, and the incentive to deviate is only privately observable. Since it is assumed that the gain from deviating can become infinitely large – that is, that the short-term gains completely dominate the long-term losses of a breakdown of the agreement – any cooperative tariff strictly lower than the Nash tariff is associated with a positive probability of cooperation breaking down in finite time, absent a safeguard provision.

It is demonstrated that by introducing an escape clause allowing for temporary deviation, cooperation can be sustained for any degree of trade liberalization. But to avoid its being applied all the time, it is necessary to attach a cost to the use of the escape clause. By making this cost a compensatory transfer to the trading partner, the use of the escape clause can be restrained and efficiency enhanced. The optimal fixed compensation cost scheme turns out to be such that the trading partner is, on average, fully compensated or, when that is not possible due to participation constraints, compensated to the largest possible extent for being exposed to deviation under the escape clause.

Apart from sustaining cooperation, there are two additional benefits from implementing an escape clause with the optimal compensation cost scheme, as compared to the case when no escape clause is included. First, expected per-period payoffs increase for any given degree of liberalization and second, liberalization can be pushed further.

The third essay, "Optimal Time Limits on Safeguards", addresses the issue of for how long countries should be allowed to withdraw commitments made under a trade agreement. The starting point is that liberalization usually creates winners as well as losers, which may be hard to identify ex ante. Ex post, some sectors will actually be worse off under liberalization and thus be in need of protection. A safeguard allowing for scaling back liberalization to protect these sectors is therefore desirable.

Protection granted to a sector in one country comes at a cost to its trading partner, however. The agreement must thus strike a balance between the benefits to the country being allowed a safeguard to be maintained for an extended period, and the cost to its trading partner from being denied market access. One way of doing this is to impose a time limit on the use of the safeguard. Negotiations preceding a trade agreement will thus, apart from determining the degree of liberalization, focus on the optimal length for applying protection under a safeguard.

In a model with two countries and an infinite number of sectors, each sector can be in either of two states at any point in time. Being in the good state, there are gains to be made from liberalization, but in the bad state losses will be incurred under liberalization and protection is therefore desirable. In each sector, switches between states are governed by independent Poisson processes. It will be shown that the optimal length of time for a safeguard to be in place depends on the ratio between the gain from liberalization under the good state and the loss from liberalization under the bad state. While no upper limit on protection is optimal for low ratios, and safeguards should thus be allowed to be applied whenever and as long as necessary to protect losers from liberalization, not including any safeguard at all for high ratios is optimal. For intermediate ratios, it is, however, optimal to have a safeguard, albeit with a finite limit on the duration of its use.

An ex ante agreed-upon time limit on the use of the safeguard will ex post be increasingly suboptimal, however. When the agreement is negotiated, the prospect of needing to apply the safeguard in excess of its time limit lies in the distant future and hence, a low weight will be attributed to this possibility. Since the number of sectors having been in the bad state for a length of time exceeding the time limit will increase monotonously over time, the maximum duration for applying the safeguard will increasingly be perceived as too short. Hence, the dissatisfaction with the agreement will grow over time.

The model just presented assumes the two countries to be completely symmetric. To understand the effect of asymmetry across countries, a modified version of the model, where only one of the countries is exposed to stochastic switches between states, is considered. The solution yielded through Nash bargaining turns out to be qualitatively similar to that obtained in the symmetric case. The solutions derived in both the symmetric and the asymmetric case may not be globally optimal, however. Hence, by further adding flexibility to the safeguard regimes, global efficiency might be enhanced.

The overall focus of the thesis is on the theoretical foundations for safeguards. Let me conclude by, admittedly very bravely, trying to make some inferences concerning the appropriate design of actual safeguards. The first essay does not go further in this respect than to point to the need for a safeguard instrument, by highlighting the limits of what countries can achieve absent safeguards. The second and the third essays then explore two different reasons for safeguards. In practice, these reasons are likely to exist simultaneously, and it is therefore necessary to have a safeguard that can address both types of situations. Building on the results above, some conclusions can be drawn on how such a provision might be structured. Whereas in the second essay the safeguard is introduced as a means of countering incentives to deviate that are only privately observable, the safeguard of the last essay serves to alleviate verifiable losses from liberalization that may be incurred in some sectors. The safeguard provision should thus preferably be able to deal with both politically motivated and unobservable incentives to deviate from liberalization commitments, and the exposure to observable economic losses from liberalization. The difference is, however, that in the former case, where information is private, there is a moral hazard problem which does not exist in the latter case, and which is countered by the requirement to compensate the trading partner for the use of the safeguard. This seems to suggest that the design of a safeguard could be along the following lines: if a verifiable loss stemming from liberalization can be established, a country shall be free to use the safeguard for a pre-specified amount of time. But if the underlying reasons for a country wishing to apply the safeguard cannot be verified, while being permitted to implement the safeguard, it should compensate the trading partner for being exposed to the safeguard.

Finally, if a safeguard of the suggested type were to be included in a trade agreement, it would most likely require a third party dispute settlement mechanism. Obviously, the implementation of such a design requires a means of agreeing on whether the underlying reasons for invoking the safeguard is of one type or the other, to determine whether compensation should be paid. It would therefore be necessary to establish a supranational agency serving to verify the underlying reasons for applying a safeguard. The dispute settlement mechanism of the World Trade Organization, for instance, seems to have played this role, having adjudicated a number of disputes involving contested safeguard measures. Its more precise role is, however, beyond the scope of this dissertation.

Chapter 2

Can Self-Destructive Trade Agreements Be Optimal?*

1 Introduction

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

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

* I am grateful to my advisors Harry Flam and Henrik Horn. Thanks also to Christina Lönnblad for editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

by threatening to punish current-period deviations in future periods, the incentive to deviate can be balanced and cooperation be sustained forever. Most models assume an infinite repetition of the one-shot Prisoner's Dilemma game, but a finite number of periods is sufficient for cooperation to be established if there exist multiple Nash equilibria, as demonstrated by Dixit (1987).

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

Although a key assumption in the literature of repeated Prisoner's Dilemma games is that a country cannot observe its trading partner's choice of tariff in the current period, and thus does not know whether it cooperates or deviates, it is often assumed that there is certainty concerning the environment where the trading partner takes its decision. More specifically, the short-term gain of deviating as well as the long-term loss of being punished for a deviation are common knowledge.

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

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

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

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

2 Uncertainty and Strategic Interaction

Introducing some sort of uncertainty into the conditions, under which decisions of complying with or breaching commitments made under a trade agreement are taken not only makes the analysis more complicated, but may also lead to different implications for the prospects of sustaining a cooperative arrangement. Several attempts have been made to incorporate uncertainty into the Prisoner's Dilemma setting of trade agreements. Hardly surprising, it is easy to find close correspondences to the literature on collusion under uncertainty. When reviewing the most important contributions within the field of uncertainty and strategic interaction, it is necessary to distinguish between different types of uncertainty. There are, broadly, three categories of uncertainty that have been addressed in the industrial organization (IO) literature and, to a lesser extent, also in the literature on trade agreements.

2.1 Ex ante Uncertainty About Commonly Observed Shocks

Strategically interacting parties may be subject to ex ante uncertainty regarding a commonly observed shock which has an impact on the incentive to deviate from a cooperative arrangement. This type of uncertainty was first addressed by Rotemberg and Saloner (1986). In their model, two competing firms are subject to ex ante unknown fluctuations in demand which, in turn, lead to fluctuating incentives to deviate. In this setting, a cooperative arrangement between the two firms will prescribe cooperation in periods of low demand, when the incentive to deviate is not sufficiently strong to make deviation worthwhile, and deviation in periods of high demand, when deviation is preferred to cooperation. Alternatively, it is possible to let the extent of cooperation vary with the realizations of demand. Crucial for this model is that the realization of the uncertainty variable which has an impact on demand is perfectly observed by both parties.

In Bagwell and Staiger (1990), the case of negotiating an agreement when there is ex ante uncertainty about the incentive to deviate from an agreed-upon tariff level in a future period is considered. In a one-sector two-country partial equilibrium model, periods of high trade volumes are associated with stronger incentives to deviate so as to make terms-of-trade gains. Hence, a cooperative agreement will have to allow for the cooperative tariff to adjust in order to dampen trade volume fluctuations and hence, counter the incentive to deviate. This type of trade management can thus be seen as an attempt by countries to maintain the self-enforcing nature of existing international cooperation. The setting and the results in Bagwell and Staiger (1990) are similar to the results in Rotemberg and Saloner (1986). In periods of high trade volumes (demand), the incentive to deviate increases and thus, cooperation will be at a lower level.

2.2 Hidden-action-type of Uncertainty

The second category of uncertainty concerns the unobservability of the strategic partner's action. As mentioned before, it is an underlying assumption in all repeated Prisoner's Dilemma models that the action taken by the opponent is not observable in the current period. In the following period, however, the choice taken by the opponent becomes common knowledge, either directly, or indirectly through inference. In this category of uncertainty models it is, however, assumed that even ex post it is not possible to verify or correctly infer the action taken by the opponent in the previous period, because the commonly observed outcome is not only influenced by the actions taken, but also by some stochastic variable. A worse-than-expected outcome can thus be the consequence of either deviation on behalf of the opponent, or a bad realization of the stochastic variable.

Green and Porter (1984), the first paper to describe strategic interaction under this type of uncertainty, investigates how collusion in a Cournot duopoly is affected by unobservable demand shocks influencing the price level. In this model, a single firm will not know whether a low market price is due to cheating by the competitor or low demand. By prespecifying a certain price level below which reversion to the Cournot Nash equilibrium will take place, it is possible to achieve collusive behavior. Episodes of Cournot Nash reversion, following low price realizations, will nevertheless occur, in order to sustain the agreement. Abreu, Pearce and Stacchetti (1986) modify and generalize the Green-Porter model and show that in equilibrium, only two quantities are ever produced. It is also shown that a firm simply needs to remember the price in the previous period and what quantity was specified by the equilibrium in that period.

The effect of this type of uncertainty on the sustainability of trade agreements was first examined by Riezman (1991).[1] Starting out from the Dixit (1987) model, a random component attached to home imports, reflecting shocks to preferences or endowments, is introduced, and it is assumed that protection is not perfectly observable.[2] This assumption is analogous to the assumption in Green and Porter (1984) and Abreu, Pearce and Stacchetti (1986), that a firm's output level cannot be observed by its competitor. When import trigger strategies are applied, reversionary

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

[2] When protection is perfectly observable, the results are similar to those obtained in Bagwell and Staiger (1990).

(high tariff) episodes are triggered by the random variable. These Nash reversions are not the result of deviation against the low-tariff agreement, but necessary to provide the incentives for sustaining cooperative low tariff episodes. This result corresponds to that obtained in Green and Porter (1984). Low (high) realizations of the price (imports) trigger reversionary episodes. If terms of trade trigger strategies are used, however, sustaining cooperation is no longer possible since, in contrast to the oligopoly case, countries have opposite incentives to influence the strategic variable and hence, any change in the terms of trade will trigger Nash reversion. In this case, cooperation can only be attained through asymmetric punishments.

2.3 Hidden-information-type of Uncertainty

The final category of uncertainty concerns hidden information. In this case, each party is exposed to a random variable, the realization of which cannot be observed by the opponent. Three subcategories within this literature can be identified, according to the type of strategic interaction. First, there are models implicitly assuming cooperative behavior among the players. This analysis has focused on the potential role for information sharing and signalling. Second, the impact of hidden information on non-cooperative equilibria has been explored. While there exist several game theoretic and oligopoly models in these two subcategories[3] , only the trade-policy related papers will be discussed here. The third and final subcategory, which has only recently been addressed in both the IO and the trade literature, concerns the impact of hidden information on cooperation in non-cooperative games. The present paper conceptually belongs to this subcategory.

2.3.1 Cooperative Games

With regard to hidden-information type of uncertainty in the context of trade cooperation, the case of one-stage games with implicitly assumed cooperative behavior has been examined by Feenstra (1987) and Feenstra and Lewis (1991). Given the underlying assumption that countries prefer to cooperate and apply agreed-upon

[3] See, for example, Vives (1984), Cramton and Palfrey (1990), Spulber (1995), and Kandori and Matsushima (1998).

policies, the problem addressed is how hidden information creates incentives to misrepresent in order to make gains.

The starting point for Feenstra (1987) is that limitations on the use of temporary import restrictions under Article XIX of the General Agreement on Tariffs and Trade (GATT) may be ineffective when one country cannot actually verify the conditions faced by an industry in the other country. In fact, incomplete information may create incentives to misrepresent the conditions faced by domestic industries in order to obtain protection.[4] In a two-country two-good model, it is assumed that production possibilities in the home country can be in either of two states, the true state being unobservable to the trading partner. Under full information, the first-best equilibrium in each state can be restored through import tariffs and equal export subsidies in the other country, whereby incomes are transferred across countries which can, of course, also be achieved by explicit transfers. If the true state in the home country cannot be observed abroad, the home country has the possibility of misrepresentation. Because of the uncertainty on behalf of the trading partner, a state-contingent agreement is no longer feasible. Incentive compatibility constraints, which are necessary to ensure the truthful revelation of the state by the home country, may then yield second-best solutions only. In an extension, the game is repeated. In this case, transfers can be intertemporal, something that may introduce new incentive compatibility constraints, however.

Feenstra and Lewis (1991) address the problem of bilateral bargaining under asymmetric information, where bargaining is conducted in a cooperative fashion, e.g. under the auspices of the GATT. Commitment to cooperative behavior is thus implicitly assumed. Informational asymmetry arises, since the home government is exposed to political pressure to restrict trade, something that cannot be directly observed by the foreign government. Subject to the constraints of incentive compatibility and both countries becoming no worse off, the optimal trade policies are shown to be tariff quotas, where a tariff is applied to imports exceeding some quota limit. By letting this limit vary, revenues and rents are allocated between the two countries in a way ensuring the truthful revelation of political pressures in the home country.

[4] Another possibility, addressed by Mitchell and Mori (2004), is informational asymmetry between industries and government within a country, which may give rise to misrepresentation by the former vis-à-vis the latter.

Thus, transferring rents from trade restrictions can be regarded as having an informational role. The incentives to apply trade restrictions are offset by compensating the trading partner through these rents and hence, domestic political pressures are truthfully reported. Such transfers are similar both to the use of voluntary export restraints and the case with safeguards and monetary compensation.

2.3.2 Non-cooperative Outcomes in Non-cooperative Games

While Feenstra (1987) and Feenstra and Lewis (1991) focus on the effect of hidden information on outcomes in cooperative games, Jensen and Thursby (1990) investigate the effect of private information on noncooperative equilibria and how the incentives for governments to establish tariff reputations might be influenced. In a two-country model, the home government is assumed to have private information about its tariff reaction function, making it either a low- or high-tariff type. It is shown that in a one-shot game, the foreign country's Nash equilibrium tariff will be lower (higher) than when it knows the home country's government to be a low- (high-) tariff type. In a two-stage game setting, a low-type home government may have incentives to establish a reputation as a high type at the initial stage, whereby the foreign government will set a lower tariff at that stage. This will be the case if the discount factor is sufficiently high, since the first-period loss due to misrepresentation is then outweighed by the potential gain of establishing a reputation inducing lower foreign tariffs.

2.3.3 Cooperative Outcomes in Non-cooperative Games

Athey and Bagwell (2001) identify two important obstacles to collusion between firms, impatience and private information. To elaborate on the impact of hidden information, an infinitely repeated Bertrand game between two firms with publicly observed prices and inelastic demand is set up. In every period, each firm is exposed to a privately observed, identically and independently distributed (i.i.d.) cost shock, making unit-cost realizations either high or low. In static frameworks with private cost information, efficient production allocation between cartel members requires communication and transfers, the latter to create the right incentives for the former to be truthful. In the base case of the infinitely repeated setting, it is assumed that while communication is possible, side payments are illegal. However, by factoring the perfect public equilibrium firm payoffs into current-period payoffs and discounted continuation values, the latter can be regarded as playing the role of side-payments, whereby a self-enforcing arrangement is possible. More specifically, instead of having a side payment from one firm to the other, one firm is favored over the other through future market-share favors exchanged in periods where both firms are equally efficient. It is shown that for a sufficiently high discount factor, first-best profits can be achieved in every period. With regard to the role of communication, both benefits and costs are identified. Benefits arise due to the fact that communication allows firms to smoothly divide the market on a state-contingent basis, while costs are incurred due to increased incentives to undercut prescribed prices. Still, it is demonstrated that in the absence of communication, first-best profits can be achieved if the discount factor is sufficiently large. Finally, the assumption of no side payments is relaxed. It is shown that, as a means of transferring payoffs, future market-share favors are strictly preferred to side payments, unless the latter are perfectly efficient (i.e. detection is of no concern).

Athey, Bagwell and Sanchirico (2004) employ an infinitely repeated Bertrand game between n firms. While costs are privately observed, prices are publicly observed before the beginning of the next period. For each firm, there is a continuum of unit cost levels, the realizations of which are i.i.d. across firms and time. In contrast to Athey and Bagwell (2001) who apply the asymmetric perfect public equilibrium concept, symmetric perfect public equilibria are derived. These prescribe a price for each cost type and an associated equilibrium continuation value for each vector of current prices in any period, the continuation value being symmetric across firms. In the unique Nash equilibrium of the stage game, the symmetric pricing strategy is strictly increasing in the firm's cost level. Such a fully sorting (strictly increasing) pricing scheme yields an efficiency benefit, because sales are allocated to the firm with the lowest costs. Considering the infinite repetition of the stage game, it is shown that fully sorting collusive schemes will not make firms better off than under the Nash-pricing scheme. If the full class of symmetric perfect public equilibria collusion schemes is considered, however, an optimal scheme can be achieved without recourse to equilibrium-path price wars, when the discount factor is sufficiently large. Moreover, a rigid-pricing scheme, where all firms set the price equal to the consumer's reservation price in every period, will result if the distribution of cost types is log-concave. Even when the discount factor is not sufficiently large to enforce the rigid-price scheme, it is still possible to adopt a partially rigid scheme where the price of lower-cost types is reduced to mitigate the incentive to cheat.

The starting point for Martin and Vergote (2004) is that the underlying reasons for antidumping are strategic in nature, something that can be supported by evidence suggesting that a significant motive behind antidumping filings lies in its retaliatory use by the parties involved. In a repeated two-good two-country setting with unobserved political preferences, antidumping is used to let tariffs adjust to changes in these preferences. It is shown that the terms-of-trade gains that can be made by increasing tariffs beyond the efficiency levels are offset in the presence of transfers or export subsidies. In the absence of these instruments, where the former is rarely used and the latter is restricted under the GATT and the World Trade Organization (WTO), truthfulness about political preferences can be achieved by letting present actions influence expected future payoffs through retaliation, and the retaliatory use of antidumping may then improve welfare if static rules governing its use are adopted. These results suggest that when the use of some instruments are restricted, the strategic or retaliatory use of the remaining ones, such as antidumping, may be the most efficient way of dealing with hidden information.

While the present paper also attempts to shed light on the impact of hidden information on cooperative outcomes in an infinitely repeated non-cooperative setting, its main focus is the trade-off between the degree and the sustainability of cooperation. The starting point is an agreement only specifying cooperative tariff levels and not allowing for any kind of direct or indirect transfers, neither within nor across periods. Hence, no sophisticated collusive schemes, like in Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004), are considered. Arguably, it is realistic to assume that such schemes are harder to implement and enforce in an environment susceptible to political considerations. It is also implicitly assumed that no other trade-distorting instrument than tariffs is available. Hence, the agreement does not include any safeguard provisions, like the antidumping instrument in Martin and Vergote (2004).[5] A government facing strong incentives to deviate will thus have no other choice than raising tariffs, thereby breaching the agreement. Naturally, this is a major simplification, but it is made to gain insights into what degree of liberalization will be chosen ex ante, when ex post gains from deviation may threaten the sustainability of the agreement.

3 The Model

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

A simpler and analytically more tractable way of investigating the impact of uncertainty on the conditions for strategic interaction and cooperation in an infinitely repeated setting is to let the government weigh current-period and future-period payoffs and let one of these weights (and hence, the relative weight) be randomly determined. Intuitively, attributing more weight to profits in the import-competing sector should be similar to giving more weight to present vis-à-vis future payoffs. In both cases the incentives for protection increase. In fact, as shown in Appendix B, these two approaches are similar. But to establish a trade-off between liberalization and sustainability of cooperation in the presence of hidden information, applying a modified PROF is analytically more difficult than the approach chosen here.

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

3.1 Introducing Discount Rate Uncertainty

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

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

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

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

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

There exists a best-response function $t_D(t^*) \equiv \arg\max_t w(t, t^*)$. From the additive separability of w, it immediately follows that the within-period best-reply tariff t_D is independent of t^*. While w increases in t for $t < t_D$ and decreases in t for $t > t_D$ (as long as trade takes place), w falls monotonously in t^* (as long as trade takes place).

In the absence of cooperation, both countries apply the optimal tariff vis-à-vis each other, i.e. $t = t^* = t_D$, and both receive the Nash equilibrium current-period payoff $w_N = w(t_D, t_D)$. Since the current-period payoff in every future period will be equal to w_N, the government's payoff in the absence of cooperation will be given by $W_N^\delta = \delta w_N + (1-\delta)w_N = w_N$, i.e. it will be independent of δ.

Before the infinitely repeated game is played, the two countries may agree to implement a cooperative tariff $t_C < t_N$ and agree on how deviations should be punished. Once the game starts, each country will choose between applying the agreed-upon cooperative tariff and the optimal tariff t_D vis-à-vis the other country. Actually, a country can choose its tariff level from a continuum. However, assuming that setting the tariff different from t_C is regarded as a deviation, a country's choice will, in fact, be binary, i.e. between applying t_C and t_D.

The current-period payoff under mutual cooperation is given by $w_C = w(t_C, t_C)$. If one country decides to break its commitment by applying the optimal tariff vis-à-vis its trading partner, it gets the current-period payoff $w_D = w(t_D, t_C)$, while its trading partner receives the sucker's payoff $w_S = w(t_C, t_D)$. If both countries apply t_D, their current-period payoff will be given by w_N. Thus, the chosen cooperative tariff level t_C does not only directly define current-period payoff under cooperation (w_C), but also indirectly defines current-period payoffs of deviation (w_D) and being deviated-against (w_S). It is straightforward that $w_C = w_D = w_S = w_N$ for $t_C = t_D$.

The following assumptions regarding the properties of $w(t, t^*)$, reflecting features of typical trade models, will be made. The payoff under mutual cooperation w_C is concave in t_C and has a unique maximum for $t_C = t_C^{opt} < t_D$. It immediately follows that there exists a $t' < t_C^{opt}$ such that $w_C = w_N$ for $t_C = t'$. Thus, $w_C > w_N$ if and only if $t' < t_C < t_D$. While the sucker's payoff w_S is concave in t_C, attaining a maximum for $t_C = t_D$, the deviator's payoff w_D decreases unambiguously in t_C. The decrease in w_D is equal to the decrease in w_C (i.e. w_D and w_C are tangent) at $t_C = t_D$ and unambiguously larger for $t_C < t_D$. Thus, $w_D - w_C$ increases, and it does so at an increasing rate as t_C decreases.

Define

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

It is easily shown that $\lim_{t_C \to t_D} \tau = 0$ and that τ increases monotonously as t_C decreases below t_D. Hence, τ can be seen as a measure of trade liberalization.

A low value of t_C corresponds to a high value of τ and thus, a high degree of trade liberalization. Since $w_C^{opt} > w_N$, the optimal degree of liberalization τ_{opt}, corresponding to $t_C = t_C^{opt}$, is strictly smaller than unity and, because $w_C = w_N$ for $t_C = t'$, the degree of liberalization τ', corresponding to $t_C = t'$, equals unity.

The relevant range of cooperative tariffs to consider is given by (t', t_D), corresponding to degrees of liberalization in the range $(0, 1)$. In this range, it is the case that $w_D > w_C > w_N > w_S$. The current-period payoff matrix is thus of Prisoner's Dilemma type.

	Cooperate	Deviate
Cooperate	w_C, w_C	w_S, w_D
Deviate	w_D, w_S	w_N, w_N

3.2 Cooperative Outcomes under Certainty

In the one-shot Prisoner's Dilemma game, both countries will choose to deviate. In fact, mutual deviation is the only equilibrium outcome. Under an infinite horizon, it is possible to create cooperation, however.[6] By threatening to punish deviations and thus associate the one-period gain from deviating with a future loss, it is possible to induce cooperative behavior. There exist many different ways of conceiving punishment phases. One way is to revert to the Nash equilibrium for a finite number of periods, before returning to the cooperative regime. It is easily shown that the longer Nash reversion lasts, the higher is the degree of liberalization that can be sustained. Hence, a grim-trigger strategy – the threat of infinite Nash reversion in case of deviation – yields the largest scope for liberalization.

In the absence of uncertainty, $\delta = \overline{\delta}$ and, applying a grim-trigger strategy and assuming a propensity for cooperative behavior across countries[7] , choosing coop-

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

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

eration yields $W_C = \overline{\delta} w_C + (1-\overline{\delta})w_C = w_C$, while opting for deviation yields $W_D = \overline{\delta} w_D + (1-\overline{\delta})w_N$. Cooperation is thus sustainable, if and only if

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

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

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

Equation (1') tells us that in order to sustain cooperation, t_C can only be reduced to the degree where τ does not exceed the discount factor $1-\overline{\delta}$. A lower $\overline{\delta}$, i.e. a higher discount factor, implies that the upper bound for liberalization $\tau_{\max}$ increases, and it is thus possible to sustain a lower t_C. The restriction given by (1') is incorporated in the negotiations concerning the cooperative tariff level and thus, imposes an upper limit on the scope for liberalization.

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

4 The Scope for Trade Liberalization under Hidden-Information-Type of Uncertainty

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

This also applies in the absence of uncertainty. However, consistent solutions are much easier to derive under certainty. Any prior regarding the likelihood of cooperative behavior of the trading partner either results in deviation or, possibly (i.e. for sufficiently low degrees of liberalization), cooperation being a consistent solution.[8] Hence, cooperation over the infinite horizon is only a matter of establishing trust, such that both countries coordinate on the cooperative solution.

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

In this section, it will first be investigated for which degrees of liberalization cooperation can be sustained for any realization of δ in the form of a self-enforcing agreement. Then, it will be explored whether there exist degrees of liberalization such that cooperation is chosen when realizations of δ are sufficiently small (i.e. when governments are sufficiently patient), while deviation is chosen when realizations of δ are high (i.e. when governments are short-termistic), in which case the ex ante likelihood of cooperation breaking down would be strictly smaller than unity, but still strictly positive.

As above in the case of certainty and throughout the analysis, it will be assumed that grim-trigger strategies are applied. From a contract theory point-of-view, it would naturally be of great interest to explore what type of punishment strategies would yield the highest expected payoff in the presence of uncertainty. It is, for

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

example, possible to consider Nash reversion for a pre-specified finite number of periods. Since this would further complicate the analysis, only the special case of grim-trigger strategies will be examined. Hence, it will be assumed that any deviation will be followed by reversion to the suboptimal Nash equilibrium forever.[9]

Let v_D and v_C be the continuation values of the game if at least one country deviates and if both countries cooperate, respectively. For any strictly positive degree of liberalization, the gain from deviating, denoted by Γ, is thus given by

$$\begin{aligned}\Gamma &= p\{[\delta w_D + (1-\delta)v_D] - [\delta w_C + (1-\delta)v_C]\} \\ &\quad + (1-p)\{[\delta w_N + (1-\delta)v_D] - [\delta w_S + (1-\delta)v_D]\} \\ &= p[\delta(w_D - w_C) + (1-\delta)(v_D - v_C)] + (1-p)\delta(w_N - w_S).\end{aligned}$$

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

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

The condition for when cooperation is chosen is thus given by

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

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

[9] Whether the threat is credible in the sense of renegotiation-proofness is obviously important to examine. This is beyond the scope of the present paper, however.

it.

That consistent beliefs are stable over time is a crucial assumption, however. The possibility of exogenous distortions to the belief formation process, i.e. what prior p is initially chosen in every period, is thus excluded. For example, it is not possible that in some future period an atmosphere of distrust suddenly emerges, making any cooperative behavior impossible.[10] In reality, this might indeed occur. A new government in one country may raise expectations of its breaching the agreement which, in turn, might induce its trading partner to deviate (pre-emptively). Such considerations are, however, assumed to be captured in the distribution of δ, which is assumed to be constant over time, whereby any derived consistent solution in the current period will also be obtained in all future periods (as long as cooperation is sustained).

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

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

It immediately follows that under a strictly positive degree of liberalization, $f(\eta) = 0$ for $\eta \leq \delta_{\min}$ and $f(\eta) = \frac{v_C - v_D}{w_D - w_C + v_C - v_D}$ for $\eta \geq \delta_{\max}$. Next, it is shown that $\eta_0 \equiv 0$ always solves equation (2).

Lemma 1 *Choosing deviation irrespective of the realization of δ is always a consistent solution for any strictly positive degree of liberalization.*

Proof. For any $\tau > 0$, which implies $w_D - w_C > 0$, $\eta = 0$, implying $\Phi(\eta) = 0$, solves equation (2). ■

Naturally, this is also true in the case of no uncertainty. It is important to emphasize that the sustainability of cooperation in infinitely repeated Prisoner's

[10] As shown below in lemma 1, choosing deviation independent of the realization of δ is an equilibrium outcome.

Dilemma games does not only depend on the discount factor, but also on the prior regarding the likelihood of the opponent opting for cooperation. By updating the prior, a consistent best response is derived. Hence, the degree of trust in the opponent is of crucial importance for sustaining a cooperative regime. Having low faith in the opponent results in deviation being the best response, under certainty as well as under uncertainty.[11]

Whether or not, and under what conditions, there exist further solutions obviously depends on the model and the distribution function specifications. More generally, it can, however, be shown that for sufficiently low τ, there exists a $\eta \geq \delta_{\max}$, implying a probability of unity that cooperation is chosen, which solves equation (2).

Lemma 2 *Choosing cooperation irrespective of the realization of δ is a consistent solution if and only if $\tau \leq 1 - \delta_{\max}$.*

Proof. If $\eta \geq \delta_{\max}$, $\Phi(\eta) = 1$ and thus $v_C - v_D = w_C - w_N$. Equation (2) then becomes

$$\eta = \frac{w_C - w_N}{w_D - w_C + w_C - w_N} = \frac{w_C - w_N}{w_D - w_N} = 1 - \tau.$$

Hence $\eta = 1 - \tau$, associated with $\Phi(\eta) = 1$, is a consistent solution if and only if

$$\delta_{\max} \leq \eta = 1 - \tau \Leftrightarrow \tau \leq 1 - \delta_{\max}.$$

Thus $\eta = 1 - \tau$, implying that $\Phi(\eta) = 1$, is a consistent solution for $\tau \in [0, 1 - \delta_{\max}]$. ■

It should be noted that cooperation will be strictly preferred to deviation for any realization of δ when $\tau < 1 - \delta_{\max}$. When $\tau = 1 - \delta_{\max}$, cooperation will be strictly preferred for $\delta < \delta_{\max}$, while a government will be indifferent between cooperating and deviating when $\delta = \delta_{\max}$.[12]

In the absence of uncertainty cooperation is sustainable if and only if $\tau \leq 1 - \overline{\delta} = \tau_{\max}$. Therefore, it is not surprising that, under uncertainty, cooperation is always

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

[12] Note also that in the special case of $\delta_{\max} = 1$, always choosing cooperation is a consistent solution only for $\tau = 0$, in which case the effects of cooperating and deviating coincide.

only possible as long as it is the preferred choice even in the worst case, i.e. for $\delta = \delta_{\max}$.

Let $\widetilde{\tau}_{\max} \equiv 1 - \delta_{\max}$ be the maximum degree of cooperation, at which cooperation can be sustained with a probability of unity. In what follows, agreements under which cooperation can be sustained forever will be referred to as safe, while agreements under which the ex ante likelihood of cooperation being chosen is strictly positive, but also strictly smaller than one, will be referred to as self-destructive. As shown above, safe agreements exist for $\tau \leq \widetilde{\tau}_{\max}$, while self-destructive agreements may or may not exist. The following lemma demonstrates that the range of degrees of trade liberalization supporting safe agreements under uncertainty is smaller than in the absence of uncertainty.

Lemma 3 *The maximum degree of liberalization under uncertainty, at which liberalization can be sustained with a probability of unity, is strictly smaller than the maximum degree of cooperation under certainty ($\widetilde{\tau}_{\max} < \tau_{\max}$).*

Proof. Since $\delta_{\max} > \overline{\delta}$ under uncertainty, it immediately follows that $\widetilde{\tau}_{\max} = 1 - \delta_{\max} < 1 - \overline{\delta} = \tau_{\max}$. ■

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

To find out how the prospects of reaching an agreement are influenced when the degree of liberalization is pushed beyond $\widetilde{\tau}_{\max}$, it is necessary to calculate the continuation values v_D and v_C (see Appendix A for derivation)

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

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

As shown above, cooperation can be sustained in every period for $\tau \leq \widetilde{\tau}_{\max}$, and the continuation value under cooperation v_C will thus equal the payoff under cooperation, w_C. Pushing τ beyond $\widetilde{\tau}_{\max}$, it is no longer the case that v_C is solely determined by the degree of liberalization τ; it will also be a function of the threshold value η and hence the density function φ, which will further complicate equation (2). An increase in τ beyond $\widetilde{\tau}_{\max}$ will thus affect v_C in two ways, directly through τ and w_D, and indirectly through η. To establish the impact of τ on η, the expressions for v_D and v_C given by (3) and (4) must be plugged into equation (2). Rearrangements then yield the consistent solution equation (CSE) (see Appendix A for derivation)

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

Given a degree of liberalization, the CSE must hold in order to obtain a threshold value that is consistent, in the sense that its implied ex ante likelihood of cooperation being chosen equals the probability attributed to the trading partner choosing cooperation. Solutions of the CSE are given by intersections of η and $f(\eta)$. As demonstrated in the lemmas above, $\eta = 0$ solves the CSE for $\tau > 0$, while $\eta = 1 - \tau$ solves the CSE for $\tau \leq \widetilde{\tau}_{\max}$. To establish whether there exist any further solutions, it is necessary to more closely examine the properties of $f(\eta)$.

Continuity of the density function implies that $f(\eta)$ is continuous in η for $\tau > 0$. As established above, $f(\eta) = 0$ for $\eta \leq \delta_{\min}$ when $\tau > 0$, and $f(\eta) = 1 - \tau$ for $\eta \geq \delta_{\max}$. Moreover, $f(\eta) \geq 0$ for $\tau \leq \frac{\overline{\delta}}{E(\delta|\delta\leq\eta)}$. Since $\frac{\overline{\delta}}{E(\delta|\delta\leq\eta)} \geq 1$, it immediately follows that $f(\eta) \geq 0$ for $\tau \leq 1$, i.e. $f(\eta) \geq 0$ in the relevant interval. It is also easily verified that $f(\eta) \leq 1$, because $f(\eta) = 1$ for $\eta > \delta_{\min}$ when $\tau = 0$ and $f(\eta)$ unambiguously decreases in τ for any $\eta > \delta_{\min}$.

The figure below demonstrates the impact of τ on the solution(s) of the CSE. Using a symmetric density function and setting $\overline{\delta} = 0.5$, $\delta_{\min} = 0.1$ and $\delta_{\max} = 0.9$, $f(\eta)$ is plotted for different values of τ (thick lines). Consistent solutions, i.e. solutions to the CSE, are given by intersections with the upward-sloping thin line, which represents the left-hand side of the CSE. The dashed vertical line denotes $\delta_{\max}$; any intersections to its right thus imply a solution with an associated probability of cooperation being chosen equal to one.

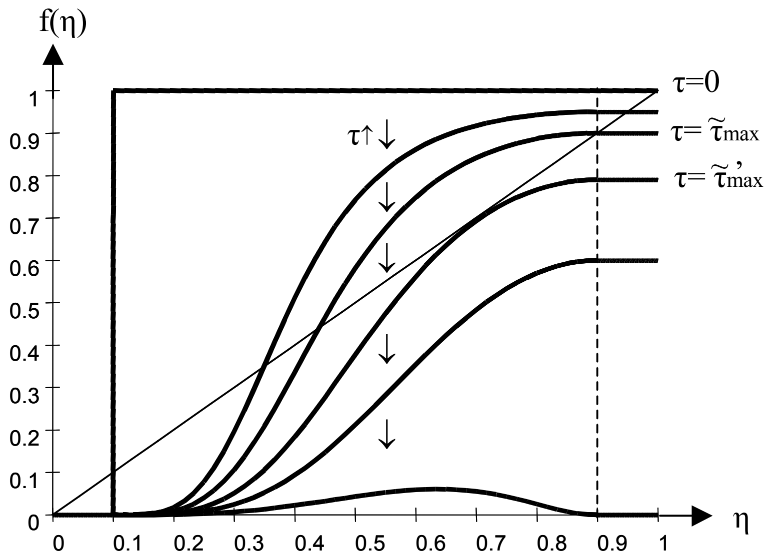

Figure 2.1 $f(\eta)$ for different τ

When the degree of liberalization is zero[13] , $f(\eta) = 0$ for $\eta \in [0, \delta_{\min}]$ and $f(\eta) = 1$ for all $\eta \in (\delta_{\min}, 1]$. Hence, $\eta_0 = 0$ and $\eta = 1$ are consistent solutions. Letting τ increase above zero will lead to a decrease in $f(\eta)$ for all $\eta \in (\delta_{\min}, 1]$. Hence, $f(\eta) < 1$, but there exist $\eta \in (0, 1)$, for which $f(\eta) > \eta$. Thus, there will be at least two intersections between η and $f(\eta)$ in this interval. Let $\eta_1 \equiv \max\{\eta \in [0, 1] | \eta = f(\eta)\}$ be the largest and $\eta_2 \equiv \min\{\eta \in (0, 1] | \eta = f(\eta)\}$ the smallest strictly positive solutions to the CSE, respectively. An increase in τ will reduce $f(\eta)$ for all $\eta \in (\delta_{\min}, 1]$ and thus, lead to a decrease in η_1 and an increase in η_2. As long as $\tau \leq \widetilde{\tau}_{\max} = 1 - \delta_{\max}$, $\eta_1 \geq \delta_{\max}$ and always cooperating is a consistent solution.

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

As τ is increased beyond $\widetilde{\tau}_{\max}$, η_1 falls below $\delta_{\max}$, while η_2 continues to increase.[14] Eventually, η_1 and η_2 will coincide, i.e. $f(\eta) < \eta$ for all $\eta > 0$ except $\eta = \eta_1 = \eta_2$ and $f(\eta)$ will be tangent to η for $\eta = \eta_1$. Naturally the degree of liberalization at which η_1 and η_2 coincide is $\widetilde{\tau}'_{\max}$, the highest degree of liberalization, for which the ex ante probability of choosing cooperation is strictly positive. Pushing the degree of liberalization beyond $\widetilde{\tau}'_{\max}$ will lead to $f(\eta) < \eta$ for all $\eta > 0$, i.e. $\eta_0 = 0$ will be the only consistent solution. Hence, η_2 no longer exists and η_1 coincides with η_0 for $\tau > \widetilde{\tau}'_{\max}$.

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

Lemma 4 *Among the solutions to the CSE, the largest yields the highest continuation value.*

Proof. See Appendix A. ■

Henceforth, it will be assumed that governments wish to behave as cooperatively as possible and thus apply η_1 for any given degree of liberalization, because this yields the largest continuation value among all consistent solutions. Thus, each government has an interest in fostering a belief about itself acting as cooperatively as possible, such that η_1 can be derived.[15] [16]

In the above figure, there exist strictly positive solutions when τ is pushed beyond $\widetilde{\tau}_{\max}$, i.e. $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$. But it need not be the case that self-destructive agreements are feasible. However, it is possible to derive a condition for when $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, as demonstrated by the next proposition.

---

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

[15] This assumption is also implicitly made in the case of no uncertainty. Note that conditions (1) and (1') both rest on the assumption that $p = 1$. Hence, it is taken as given that governments want to act as cooperatively as possible.

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

Proposition 1 *Assuming that there exist no more than three solutions to the CSE, self-destructive agreements exist (i.e. $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$), if and only if*

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

Proof. See Appendix A. ■

The intuition behind this result is the following. A low density for $\delta = \delta_{\max}$ implies that a threshold value η slightly smaller than $\delta_{\max}$ will be associated with a probability of cooperation being chosen negligibly smaller than unity. Hence, the effect on the likelihood of cooperation is only negligibly different from when there is no uncertainty, in which case cooperation can be sustained when τ is marginally increased beyond $\widetilde{\tau}_{\max}$, because $\tau_{\max} > \widetilde{\tau}_{\max}$. However, if $\varphi(\delta_{\max})$ is too large, this will no longer be the case. A sufficiently low density at $\delta_{\max}$ will thus ensure the existence of $\tau > \widetilde{\tau}_{\max}$ that are associated with $\eta_1(\tau) \in (\delta_{\min}, \delta_{\max})$ and hence, an implied probability of cooperation being chosen of $\Phi(\eta_1(\tau)) \in (0, 1)$.

It is important to note that the above derived condition rests on the assumption of no more than three solutions to the CSE. If it is possible that there exist more than three solutions to the CSE[17] , the condition is sufficient to ensure $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, but it is still possible that $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, even if it does not hold. A similar condition to that obtained in the previous proposition can easily be derived. If there exists at least one consistent solution $\eta > \delta_{\min}$, for which $f'(\eta) < 1$ when $\tau = \widetilde{\tau}_{\max}$ and for which the density is sufficiently low, then $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$.

Two main conclusions can be drawn from this section. First, the introduction of uncertainty will unambiguously diminish the range of degrees of liberalization where it is possible to sustain cooperation with a probability of unity. Second, uncertainty

[17] Pinning down the maximum possible number of solutions to the CSE turns out to be very difficult. Obviously, it is the density function of the random variable that determines how many such solutions exist. However, it is analytically untractable to link the properties of the density function to the shape of $f(\eta)$. It seems that the number of peaks of the density function determine how many solutions to the CSE there may exist. At most one peak (i.e. a well-behaved density function) limits the maximum number of solutions to three, two peaks are associated with a maximum number of five, and so on.

may create a range of degrees of liberalization, which is associated with a probability of cooperation occurring that is strictly positive, but strictly smaller than unity. High degrees of liberalization will, in this case, be associated with a strictly positive likelihood of deviation being chosen, and cooperation will thus break down in finite time. Under uncertainty, countries may hence face a tradeoff between the gains from further liberalization and the increased risk of cooperation breaking down.

5 The Optimal Degree of Liberalization

The previous section investigated for which degrees of liberalization safe agreements are possible ($\tau \leq \widetilde{\tau}_{\max}$) and under which conditions self-destructive agreements are feasible ($\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$). The far more important question, however, is what degree of liberalization is optimal under uncertainty. Will the optimal solution always be a degree of liberalization such that the agreement will never break down, or will it be optimal to let the agreement entail the seeds for its own eventual self-destruction? Ex ante, countries may have to decide between applying a lower degree of liberalization guaranteeing cooperative behavior on the one hand and, if that is possible, applying a higher degree of liberalization associated with a risk of breakdown of cooperation, on the other hand. In this section, it will be explored under which conditions the optimal agreement is safe or self-destructive.

5.1 The First-Best Outcome

Before addressing that issue, the best conceivable agreement will be determined. The following proposition asserts that under uncertainty, no better outcome than the best possible outcome in the absence of uncertainty can be achieved.

Proposition 2 *The best possible outcome in the presence of uncertainty is to apply τ_{opt} with an associated probability of choosing cooperation equal to unity.*

Proof. Given some strictly positive degree of liberalization, $v_C = w_N$ for consistent solutions $\eta \leq \delta_{\min}$ and $v_C = w_C$ for consistent solutions $\eta \geq \delta_{\max}$. From lemma 4, we can infer that $\frac{dv_C}{d\eta} > 0$ for any consistent solutions $\eta \in (\delta_{\min}, \delta_{\max}]$. Hence, the best possible outcome must be associated with a consistent solution $\eta \geq \delta_{\max}$ and thus, a probability of unity of cooperation being chosen. Since $v_C = w_C$ for $\eta \geq \delta_{\max}$, it immediately follows that $\tau = \tau_{opt}$ maximizes v_C when $\eta \geq \delta_{\max}$. Therefore, the best possible outcome under uncertainty is applying τ_{opt} with an associated probability of choosing cooperation equal to unity. ■

Note that this proposition only states what is the conceivably most optimal outcome. Whether it is attainable depends on the CSE. Intuitively, this result seems trivial. How could, for example, a degree of liberalization larger than τ_{opt}, associated with a probability of cooperation being chosen strictly smaller than one, yield a higher continuation value? Theoretically, this could actually be the case, if the gain from deviating were very large, while the loss from being deviated against were small, for $\tau > \tau_{opt}$. It could then be ex ante desirable that cooperation would actually break down in finite time, if the (appropriately weighted) gain from deviating were sufficiently large to outweigh the (appropriately weighted) loss from being deviated against and the (appropriately weighted) loss from having a lower w_C as long as cooperation is sustained as well as the loss from a strictly positive likelihood of Nash reversion occurring in every period. However, the above proposition shows that this can never be the case.[18]

Herzing (2005) uses a similar setting and introduces a safeguard, allowing a country to deviate while compensating its trading partner. It is shown that under optimal fixed monetary compensation, it is actually optimal to increase liberalization beyond τ_{opt}. Hence, allowing for transfers across countries, the presence of hidden-information type of uncertainty might actually lead to a degree of liberalization higher than in the absence of uncertainty. In the absence of such transfers or any other flexibility-enhancing instrument in a trade agreement, this cannot be the case in the present model, however.

What can immediately be inferred from proposition 2 is that if there exists a consistent solution $\eta \geq \delta_{\max}$ for $\tau = \tau_{opt}$, then τ_{opt} should be applied. Let $\widetilde{\tau}_{opt}$ be the optimal degree of liberalization in the presence of uncertainty. The following propo-

[18] It is likely that the assumed additive separability of the government's payoff function, which ensures that $w_D - w_C = w_N - w_S$, is crucial for this result.

sition demonstrates under what conditions the optimal solution under certainty can be replicated under uncertainty.

Proposition 3 *If $\overline{\delta} < 1 - \tau_{opt}$ and $\delta_{\max} \in (\overline{\delta}, 1 - \tau_{opt}]$, then $\widetilde{\tau}_{opt} = \tau_{opt}$ and cooperation can be infinitely sustained.*

Proof. For τ_{opt} to be applicable and cooperation to always be chosen, it is necessary that $\tau_{opt} \leq \widetilde{\tau}_{\max}$

$$\tau_{opt} \leq \widetilde{\tau}_{\max} = 1 - \delta_{\max} \Leftrightarrow \delta_{\max} \leq 1 - \tau_{opt}.$$

For this condition to be relevant, it is naturally necessary that $\overline{\delta} < 1 - \tau_{opt}$. Hence, for $\overline{\delta} < 1 - \tau_{opt}$ and $\delta_{\max} \in (\overline{\delta}, 1 - \tau_{opt}]$, it will be the case that $\tau_{opt} \leq \widetilde{\tau}_{\max}$, i.e. the optimal degree of liberalization can be applied while the probability of choosing cooperation equals unity. From proposition 2, it immediately follows that $\widetilde{\tau}_{opt} = \tau_{opt}$. ■

Intuitively, this result is straightforward. If governments are expected to be patient enough to make cooperation the preferred choice even in the most myopic case under τ_{opt}, then it is optimal to set $\tau = \tau_{opt}$ and hence, the outcome under certainty can be replicated.

5.2 When the First-Best Outcome Is Unattainable

The situation when it is not possible to apply τ_{opt} such that cooperation is sustainable over the infinite horizon remains to be considered. Is it optimal to restrict the degree of liberalization such that cooperation is sustainable even for the most myopic realization of δ? Or, is it ex ante optimal to choose a degree of liberalization associated with a strictly positive probability of deviation occurring in every period?

In the previous section, a condition for when liberalization can be pushed beyond $\widetilde{\tau}_{\max}$ without cooperation instantly breaking down was derived for the case when there exist three solutions to the CSE at most. If $\varphi(\delta_{\max}) < \overline{\varphi(\delta_{\max})}$, then it is possible that there exist $\tau > \widetilde{\tau}_{\max}$, such that the ex ante probability of cooperation being chosen in any period is strictly positive, albeit strictly smaller than unity. The following proposition states under what condition increasing τ beyond $\widetilde{\tau}_{\max}$ will be optimal.

Proposition 4 *When $\delta_{\max} > 1 - \tau_{opt}$, there exists a threshold value $\overline{\varphi(\delta_{\max})}' \in (0, \overline{\varphi(\delta_{\max})}]$ such that a self-destructive agreement is preferred to a safe agreement ($\widetilde{\tau}_{opt} > \widetilde{\tau}_{\max}$) if and only if*

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

Proof. See Appendix A. ■

Hence, if the marginal likelihood of $\delta_{\max}$ occurring is sufficiently small, pushing liberalization beyond $\widetilde{\tau}_{\max}$ will be worthwhile.[19] To understand the intuition behind this result, it is important to emphasize the two opposing effects at work when τ is increased beyond $\widetilde{\tau}_{\max}$. Given that $\delta_{\max} > 1 - \tau_{opt}$ and hence $\widetilde{\tau}_{\max} < \tau_{opt}$, there is an unambiguously positive effect on the outcome under mutual cooperation and thus, on the continuation value. But there is also a negative effect stemming from the increased ex ante likelihood of deviation occurring due to η falling below $\delta_{\max}$. This negative effect is directly related to the marginal likelihood of $\delta_{\max}$ occurring. The larger is this effect, i.e. the larger is $\varphi(\delta_{\max})$, the stronger will the negative impact of an increase in τ beyond $\widetilde{\tau}_{\max}$ be. Obviously, the negative effect is non-existent if $\varphi(\delta_{\max}) = 0$, and it will be worthwhile to increase τ beyond $\widetilde{\tau}_{\max}$ as long as there is a positive effect (which is the case when $\delta_{\max} > 1 - \tau_{opt}$).

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

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

$\overline{\varphi(\delta_{\max})}'$ can be obtained such that the impact of changes in $\overline{\delta}$ and $\delta_{\max}$ can be more easily assessed

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

The most myopic realization $\delta_{\max}$ enters expression (7') both directly and indirectly through $\frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}}$. First, consider the shape of $\frac{\frac{d(w_C-w_N)}{d\tau}}{w_C-w_N}$. Since w_C is concave in τ, it immediately follows that $\frac{\frac{d(w_C-w_N)}{d\tau}}{w_C-w_N}$ is strictly decreasing in τ. Because $\frac{dw_C}{d\tau}|_{\tau=0} > 0$, $\frac{\frac{d(w_C-w_N)}{d\tau}}{w_C-w_N}$ goes to infinity as τ approaches zero, while $\frac{\frac{d(w_C-w_N)}{d\tau}}{w_C-w_N}|_{\tau=\tau_{opt}} = 0$. Since $\delta_{\max} > 1-\tau_{opt}$, it follows that $\widetilde{\tau}_{\max} < \tau_{opt}$ and hence, $\frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}} > 0$. An increase in $\delta_{\max}$ leads to a lower $\widetilde{\tau}_{\max}$ and thus a higher value for $\frac{\frac{d(w_C-w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}}{(w_C-w_N)|_{\tau=\widetilde{\tau}_{\max}}}$, its value equalling zero for $\delta_{\max} = 1-\tau_{opt}$ and going to infinity as $\delta_{\max}$ approaches unity. Therefore, the first term of the denominator will unambiguously decrease when $\delta_{\max}$ is increased. The effect on the second term of the denominator is unambiguously positive if and only if $\delta_{\max} > \frac{1-\overline{\delta}}{2}$. Hence, if $\delta_{\max} \geq \frac{1-\overline{\delta}}{2}$, an increase in $\delta_{\max}$ unambiguously increases $\overline{\varphi(\delta_{\max})}'$.[20] Moreover $\overline{\varphi(1-\tau_{opt})}' = 0$, while $\lim_{\delta_{\max}\to 1}\overline{\varphi(\delta_{\max})}' = \infty$. A low value of $\delta_{\max}$ implies a low $\overline{\varphi(\delta_{\max})}'$, and hence it will be worthwhile to increase τ beyond $\widetilde{\tau}_{\max}$ only for low marginal likelihoods of $\delta_{\max}$. A high value of $\delta_{\max}$ implies a high $\overline{\varphi(\delta_{\max})}'$ and hence, increasing τ beyond $\widetilde{\tau}_{\max}$ will be worthwhile except when the marginal likelihood of $\delta_{\max}$ occurring is very high; in fact it will always be beneficial to increase τ beyond $\widetilde{\tau}_{\max}$, if $\delta_{\max} = 1$, in which case $\widetilde{\tau}_{\max} = 0$ and no safe agreement with a strictly positive degree of liberalization is feasible.

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

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

overall impact of an increase in τ beyond $\widetilde{\tau}_{\max}$ becoming negative. For $\delta_{\max} = 1$, the positive effect of increasing τ above zero is infinitely large, thus always outweighing the negative impact on the likelihood of cooperation.

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

Lemma 5 *The scope for self-destructive agreements being preferred to safe agreements, increases if*

(i) $\overline{\delta}$ *is sufficiently large; and if*

(ii) $\delta_{\max}$ *is sufficiently large; and if*

(iii) $\varphi(\delta_{\max})$ *is sufficiently small.*

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

Lemma 6 *If* $\widetilde{\tau}'_{\max} > \widetilde{\tau}_{\max}$, *then* $\widetilde{\tau}_{opt} < \widetilde{\tau}'_{\max}$.

Proof. See Appendix A. ■

Hence, when increasing τ beyond $\widetilde{\tau}_{\max}$ is optimal, pushing liberalization as far as possible will never be optimal. It is now possible to conclude this subsection and prescribe the optimal agreement when $\delta_{\max} > 1 - \tau_{opt}$.

Proposition 5 *When the first-best outcome cannot be implemented* ($\delta_{\max} > 1 - \tau_{opt}$), *then it is optimal to implement a safe agreement such that* $\widetilde{\tau}_{opt} = \widetilde{\tau}_{\max} < \tau_{opt}$ *if and only if* $\varphi(\delta_{\max}) \geq \overline{\varphi(\delta_{\max})}'$. *If, however,* $\varphi(\delta_{\max}) < \overline{\varphi(\delta_{\max})}'$, *then it is optimal to implement a self-destructing agreement such that* $\widetilde{\tau}_{opt} \in (\widetilde{\tau}_{\max}, \widetilde{\tau}'_{\max})$.

Proof. From proposition 4, it follows that increasing τ beyond $\widetilde{\tau}_{\max}$ is not worthwhile, if $\varphi(\delta_{\max}) \geq \overline{\varphi(\delta_{\max})}'$. Since $\delta_{\max} < 1 - \tau_{opt}$, it follows that $\widetilde{\tau}_{\max} < \tau_{opt}$ and hence liberalization should be as large as possible such that cooperation is always sustainable, i.e. $\widetilde{\tau}_{opt} = \widetilde{\tau}_{\max}$. If $\varphi(\delta_{\max}) < \overline{\varphi(\delta_{\max})}'$, however, then v_C increases as τ is increased beyond $\widetilde{\tau}_{\max}$ and hence, $\widetilde{\tau}_{opt} > \widetilde{\tau}_{\max}$. From the previous lemma, it immediately follows that $\widetilde{\tau}_{opt} \in (\widetilde{\tau}_{\max}, \widetilde{\tau}'_{\max})$. ■

6 Conclusions

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

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

There are, however, ways of overcoming the problem that agreements will not always be infinitely sustainable. The remaining two essays of this thesis will explicitly address safeguard provisions similar to those found in trade agreements such as the GATT, for example, which allow signatory countries to withdraw liberalization commitments under the agreement in order to protect certain overriding interests under specified conditions, thus possibly eliminating the risk of breakdown.

7 References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Martin, Alberto and Vergote, Wouter (2004), *A Case for Strategic Antidumping*, mimeo, Columbia University, New York

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

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

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

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

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

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

8 Appendix A

8.1 Derivation of v_D and v_C

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

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

8.2 Derivation of $f(\eta)$

$$\begin{aligned}
f(\eta) &= \frac{(v_C - v_D)\Phi(\eta)}{w_D - w_C + (v_C - v_D)\Phi(\eta)} \\
&\overset{(4)}{=} \frac{[\overline{\delta}\Phi(\eta) - \tau \int_{\delta_{\min}}^{\eta} \delta d\Phi]\Phi(\eta)}{\tau\{1 - \Phi(\eta)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi]\} + [\overline{\delta}\Phi(\eta) - \tau \int_{\delta_{\min}}^{\eta} \delta d\Phi]\Phi(\eta)}
\end{aligned}$$

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

8.3 Proof of Lemma 4

Since v_C is constant for $\eta \notin [\delta_{\min}, \delta_{\max}]$, it remains to be shown that, given a degree of liberalization, v_C unambiguously increases for $\eta \in [\delta_{\min}, \delta_{\max}]$. This is not straightforward from expression (4). However, by plugging equation (5) into equation (4), the CSE is incorporated such that an expression for v_C which is multiplicatively separable in τ (which enters through $w_D - w_N$) and consistent threshold values η, is obtained.

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

$$\overset{(5)}{=} \frac{\overline{\delta}\Phi(\eta) - \frac{\overline{\delta}(1-\eta)\Phi^2(\eta)}{[1-\Phi^2(\eta)]\eta + \Phi(\eta)\int_{\delta_{\min}}^{\eta} \delta d\Phi}\int_{\delta_{\min}}^{\eta} \delta d\Phi}{1 - \Phi(\eta)[\Phi(\eta) - \int_{\delta_{\min}}^{\eta} \delta d\Phi]}(w_D - w_N)$$

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

Note that this expression is valid only for $\eta \in [\delta_{\min}, \delta_{\max}]$. For any given $\tau > 0$ (implying $w_D - w_N > 0$), the first derivative with respect to η is given by

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

It is strictly positive for $\eta > \delta_{\min}$. Hence, if more than one solution to the CSE exist, then the largest of these will yield the largest v_C.

8.4 Proof of Proposition 1

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

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

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

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

Hence,

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

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

8.5 Proof of Proposition 4

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

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

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

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

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

Thus,

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

$$\frac{\partial v_C}{\partial \tau}|_{\tau=\widetilde{\tau}_{\max}} = \delta_{\max}\frac{d(w_D - w_N)}{d\tau}|_{\tau=\widetilde{\tau}_{\max}}.$$

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

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

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

Hence,

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

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

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

Thus,

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

Hence,

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

Thus,

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

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

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

Hence,

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

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

8.6 Proof of Lemma 6

From the proof of proposition 4, we know that

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

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

9 Appendix B

9.1 A Modified PROF

In what follows, a simplified version of the PROF proposed by Baldwin (1987) will be applied in order to demonstrate a similarity to the approach taken in this paper, with regard to cooperative arrangements in the presence of hidden information.

There are two countries, each with an exporting and an import-competing sector which repeatedly interact by setting import tariffs. It suffices to consider one country only. The per-period payoff is denoted by r. Let λ, μ, ξ and σ be the weights attributed to consumer surplus CS, profits of the import-competing sector Π, tariff revenues TR and profits of the exporting sector Π^*, respectively. While consumer surplus, import-competing sector profits and tariff revenues are solely influenced by the domestic tariff t, exporting sector profits only depend on the foreign tariff t^*. Assume that ex ante, there is uncertainty regarding what weight will be given to the import-competing sector. Hence $\lambda = \overline{\lambda}$, $\xi = \overline{\xi}$ and $\sigma = \overline{\sigma}$, while $\mu \in [\mu_{\min}, \mu_{\max}]$ is stochastically determined by the density function φ_μ with an associated cumulative distribution function Φ_μ and an expected value of $\overline{\mu} \in (\mu_{\min}, \mu_{\max})$. This stochastic process is i.i.d. both across time and with regard to the corresponding stochastic process in the other country. The per-period payoff r thus depends on t and t^* as well as the realization of μ

$$r(t, t^*; \mu) = \overline{\lambda} CS(t) + \mu \Pi(t) + \overline{\xi} TR(t) + \overline{\sigma} \Pi^*(t^*).$$

In the absence of any agreement, the domestic tariff is determined as follows in every period

$$\frac{dr(t,t^*;\mu)}{dt} = \overline{\lambda}CS'(t) + \mu\Pi'(t) + \overline{\xi}TR'(t) = 0.$$

It will be assumed that there exists a unique t that solves the above first-order condition. The optimal tariff will thus be a function of the realization of μ, i.e. $t^D \equiv t^D(\mu) = \arg\max_t r(t,t^*;\mu)$. Since profits in the import-competing sector increase in the domestic tariff, i.e. $\Pi'(t) \geq 0$, the resulting optimal tariff t^D increases in μ. Hence, it will take on values in the interval $[t^D_{\min}, t^D_{\max}]$, where $t^D_{\min} = t^D(\mu_{\min})$ and $t^D_{\max} = t^D(\mu_{\max})$. In the presence of a trade agreement, this is the tariff that would be applied in case of deviation.

The two countries may agree to set tariffs below the levels that would prevail in the absence of any cooperation, in order to reap the gains from trade. The following restricting assumption will be made.

Assumption The cooperative tariff t^C is lower than any tariff that would be set in the absence of an agreement, i.e. $t^C < t^D_{\min}$.

By making this assumption, it will thus never be the case that once an agreement is in place, the optimal tariff is smaller than the agreed-upon tariff. Or, in other words, it can never be the case that there is a (short-term) incentive for a country to reduce its tariff.

Since short-term, i.e. one-period, gains can be made by deviating, the two countries find themselves in a Prisoner's Dilemma. Under an agreement, payoffs under mutual cooperation, deviation, being deviated against and mutual deviation are given by $r^C = r(t^C, t^C;\mu)$, $r^D = r(t^D(\mu), t^C;\mu)$, $r^S = r(t^C, t^D(\mu^*);\mu)$ and $r^N = r(t^D(\mu), t^D(\mu^*);\mu)$ respectively. Note that since the per-period payoff function is additively separable in components that are functions of t and t^*, respectively, it is the case that $r^D - r^C = r^N - r^S$, just as in the model applied in this paper.

The short-term gain from deviating is given by $r^D - r^C$. Its derivative with respect to t^C is given by

$$\frac{d(r^D - r^C)}{dt^C} = -\overline{\lambda}CS'(t^C) - \mu\Pi'(t^C) - \overline{\xi}TR'(t^C) = -\frac{dr(t,t^*;\mu)}{dt}|_{t=t^C}.$$

Since, by assumption $t^C < t^D(\mu)$ for any μ, it follows that $\frac{dr(t,t^*;\mu)}{dt}|_{t=t^C} > 0$ and thus $\frac{d(r^D - r^C)}{dt^C} < 0$, i.e. reducing t^C further increases the short-term gain from deviating for any given μ.

Let the continuation values for when cooperation has been sustained in the present period and for when deviation has occurred be denoted by s^C and s^D, respectively. Let the government attribute weight $\overline{\delta} \in (0,1)$ to present payoffs, the residual $1 - \overline{\delta}$ being attributed to future flows of payoffs. Let $\Omega \equiv \Omega_{t^C}(\mu)$ be the gain from deviating from cooperation, and let p be the probability that the trading partner opts for cooperation. Assuming that grim-trigger strategies are applied, it is thus the case that

$$\begin{aligned}\Omega &= p[\overline{\delta}(r^D - r^C) + (1-\overline{\delta})(s^D - s^C)] + (1-p)\overline{\delta}(r^N - r^S) \\ &= \overline{\delta}(r^D - r^C) + (1-\overline{\delta})p(s^D - s^C).\end{aligned}$$

Both continuation values solely depend on expected values and are hence independent of the current realization of μ. The impact of μ on Ω is thus given by

$$\begin{aligned}\frac{d\Omega}{d\mu} &= \overline{\delta}[\frac{dr^D}{d\mu} - \frac{dr^C}{d\mu}] \\ &= \overline{\delta}[\lambda CS'(t^D)\frac{dt^D}{d\mu} + \mu\Pi'(t^D)\frac{dt^D}{d\mu} + \Pi(t^D) + \xi TR'(t^D)\frac{dt^D}{d\mu} - \Pi(t^C)] \\ &= \overline{\delta}[\frac{dr(t,t^*)}{dt}|_{t=t^D}\frac{dt^D}{d\mu} + \Pi(t^D) - \Pi(t^C)].\end{aligned}$$

Since, by definition, $\frac{dr(t,t^*)}{dt}|_{t=t^D} = 0$ and $t^C < t^D_{\min}$, it follows that $\frac{d\Omega}{d\mu} = \Pi(t^D) - \Pi(t^C) > 0$. Hence the gain from deviating increases unambiguously in the realization of μ.

Cooperation is chosen if and only if

$$\begin{aligned}&\Omega \leq 0 \\ &\Leftrightarrow \overline{\delta}(r^D - r^C) + (1-\overline{\delta})p(s^D - s^C) \leq 0 \\ &\Leftrightarrow \frac{\overline{\delta}}{1-\overline{\delta}}(r^D - r^C) \leq p(s^C - s^D). \qquad (8)\end{aligned}$$

Since $\frac{d\Omega}{d\mu} = \frac{dr^D}{d\mu} - \frac{dr^C}{d\mu} > 0$, it immediately follows that the left-hand side of (8) is strictly increasing in μ. Since the right-hand side is constant, it is solely the realization of μ that determines whether cooperation is preferred to deviation. If $\Omega_{t^C}(\mu) \geq 0$ for all μ, then obviously $\eta_\mu \leq \mu_{\text{min}}$ and cooperation will never be chosen. And if $\Omega_{t^C}(\mu) \leq 0$ for all μ, then obviously $\eta_\mu \geq \mu_{\text{max}}$ and cooperation will always be chosen. If, however, $\Omega_{t^C}(\mu) \leq 0$ if and only if $\mu \leq \eta_\mu \in (\mu_{\text{min}}, \mu_{\text{max}})$, then the ex ante likelihood of cooperation being chosen is strictly positive, albeit strictly smaller than one. Thus, it is possible to derive a threshold value for μ, such that cooperation is chosen whenever μ is smaller than that threshold value, i.e. $\Omega \leq 0 \Leftrightarrow \mu \leq \eta_\mu$.

Assuming grim-trigger strategies to be applied, s^D equals the expected outcome under the Nash equilibrium. The continuation value s^C will, however, depend on η_μ as well as the cooperative tariff t^C. Define

$$\Psi_{t^C}(\mu) \equiv \overline{\delta}(r^D - r^C) = \overline{\delta}[r(t^D(\mu), t^C; \mu) - r(t^C, t^C; \mu)].$$

Equality in condition (8) is given by setting $\mu = \eta_\mu$. Hence,

$$\Psi_{t^C}(\eta_\mu) = (1 - \overline{\delta})p[s^C(t^C, \eta_\mu) - s^D].$$

Consistency and symmetry imply that $p = \Phi_\mu(\eta_\mu)$ must hold. Hence, consistent solutions for η_μ are implicitly determined by the following equation

$$\Psi_{t^C}(\eta_\mu) = (1 - \overline{\delta})\Phi_\mu(\eta_\mu)[s^C(t^C, \eta_\mu) - s^D]. \tag{9}$$

For any given cooperative tariff level, t^C, the left-hand side of (9) represents the short-term gain from deviating under a consistent solution η_μ, while the right-hand side of (9) expresses the expected long-term gain from sustaining cooperation under η_μ. The left-hand side of (9) unambiguously increases in η_μ and as cooperative tariffs are reduced (see above).

To conclude, the main characteristics of the incentives to deviate in the presence of hidden information can be summarized as follows:

1. The one-period gain from deviating increases as the cooperative tariff decreases.

2. The gain from opting for deviation instead of cooperation increases unam-

biguously in the current-period weight attributed to profits in the import-competing sector. Hence, higher realizations of the stochastic variable are associated with increasing incentives to deviate.

3. Given a cooperative tariff level, there exists a threshold level η_μ, such that cooperation is preferred to deviation if and only if $\mu \leq \eta_\mu$. Hence, deviation is preferred to cooperation for all realizations of the current-period weight attributed to profits in the import-competing sector (i.e. $\eta_\mu \leq \mu_{\min}$), or cooperation is preferred to deviation for all realizations of the current-period weight attributed to profits in the import-competing sector (i.e. $\eta_\mu \geq \mu_{\max}$), or cooperation is preferred to deviation only for sufficiently low realizations of μ (i.e. $\eta_\mu \in (\mu_{\min}, \mu_{\max})$).

4. Consistent solutions of the threshold value η_μ are implicitly determined through an expression equalizing short-term gains from deviating with expected long-term gains from sustaining cooperation under η_μ for any given cooperative tariff level.

9.2 Discount Rate Uncertainty Revisited

In the model applied in the present paper, the gain from deviating is given by $\Gamma = \delta(w_D - w_C) + p(1-\delta)(v_D - v_C)$. It is easily established that the current-period gain from deviating, given by $\delta(w_D - w_C)$, decreases in the cooperative tariff level, i.e. it will be larger the lower the cooperative tariff is set. This corresponds to conclusion 1 in the previous subsection.

It is easily shown that Γ unambiguously increases in δ

$$\frac{d\Gamma}{d\delta} = w_D - w_C + p(v_C - v_D) \geq 0.$$

Hence, there is also a correspondence with conclusion 2 in the previous subsection. A higher realization of the stochastic variable is unambiguously associated with an increased incentive to deviate.

Cooperation will be preferred to deviation if and only if the future gain from sustaining cooperation outweighs the current-period gain from deviating, i.e. if and only if $\Gamma \leq 0$

$$\begin{aligned}\Gamma \leq 0 &\Leftrightarrow \delta(w_D - w_C) \leq p(1-\delta)(v_C - v_D) \\ &\Leftrightarrow \frac{\delta}{1-\delta}(w_D - w_C) \leq p(v_C - v_D). \end{aligned} \tag{8'}$$

The similarity between conditions (8) and (8') is easily seen. The left-hand side of (8') unambiguously increases in δ, while the right-hand side of (8'), consisting of expected values, is constant. Hence, there exists a threshold value η such that cooperation is chosen if and only if $\delta \leq \eta$. Depending on the degree of liberalization, deviation is preferred to cooperation for all realizations of the current-period weight attributed to profits in the import-competing sector (i.e. $\eta \leq \delta_{\min}$), or cooperation is preferred to deviation for all realizations of the current-period weight attributed to profits in the import-competing sector (i.e. $\eta \geq \delta_{\max}$), or cooperation is preferred to deviation only for sufficiently low realizations of δ (i.e. $\eta \in (\delta_{\min}, \delta_{\max})$). Thus, a correspondence to conclusion 3 in the previous subsection can also be obtained.

Since grim-trigger strategies are applied, v_D equals the expected outcome under the Nash equilibrium. The continuation value v_C is, however, a function of η as well as the degree of liberalization τ. Define

$$\Delta_\tau(\delta) \equiv \delta(r^D - r^C) = \delta[r(t^D(\mu), t^C; \mu) - r(t^C, t^C; \mu)].$$

Equality in condition (8) is given by setting $\delta = \eta$. Hence,

$$\Delta_\tau(\eta) = (1-\eta)p[v_C(\tau, \eta) - v_D].$$

Consistency and symmetry imply that $p = \Phi_\mu(\eta_\mu)$ must hold. Hence, consistent solutions for η_μ are thus implicitly determined by the following equation

$$\Delta_\tau(\eta) = (1-\eta)\Phi(\eta)[v_C(\tau, \eta) - v_D]. \tag{9'}$$

For any given degree of liberalization, τ, the left-hand side of (9') represents the short-term gain from deviating under a consistent solution η, while the right-hand side of (9') expresses the expected long-term gain from sustaining cooperation under

η.[21] By comparing equations (9) and (9'), it is easy to establish a correspondence to conclusion 4 in the previous subsection. The left-hand side of (9') unambiguously increases in η and as cooperative tariffs are reduced (i.e. as τ increases).

Hence, in the present context of investigating the scope for cooperation in an infinitely repeated setting, where two governments strategically interact in the presence of hidden information due to i.i.d. (both across countries and across periods), stochastic processes with an impact on the incentives to deviate from a cooperative agreement, a modified PROF will yield similar characteristics as the objective function applied in this paper. While strict equivalence cannot be established between these two models, it is, however, possible to demonstrate that there exists a striking similarity between them in terms of the obtained expressions and conditions.

[21] Note that equation (9') is equivalent to the CSE.

Chapter 3

Escape and Optimal Compensation in Trade Agreements *

1 Introduction

Two main features of post-war trade have been a reduction of tariff levels, and an increasing use of non-tariff barriers (NTBs). As liberalization has continued and trade has expanded, countries have become increasingly exposed to foreign markets and the world market. Despite beneficial effects through increased exploitation of gains from trade, this increasing exposure to trade flows has created increased incentives for countries to temporarily deviate from cooperation to respond to political economic strains due to unforeseen circumstances. Having committed to lower tariff levels, countries have therefore increasingly made use of various contingent protection NTBs in response to domestic political economic pressures. Since the increasing use of such trade distorting measures have threatened to undermine the achievements of trade liberalization, trade negotiations have increasingly come to fo-

* I am grateful to my advisors Harry Flam and Henrik Horn. I would also like to thank Bård Harstad, Raymond Riezman and Robert Staiger, as well as participants in the European Trade Study Group's Annual Meeting in Madrid in September 2003, seminar participants at the Research Institute of Industrial Economics in Stockholm and participants in the Nordic International Trade Seminar in Copenhagen in May 2004 for helpful comments. Thanks also to Christina Lönnblad for editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

cus on removing the NTBs by finding more orderly mechanisms that allow countries to temporarily deviate from the cooperative regime under certain circumstances, without thereby threatening overall liberalization. The perception has also been that such arrangements make countries more willing to liberalize, and thus also enhances the possibilities of reaching an agreement in the first place.

In the present paper, two symmetric countries strategically interact in an infinitely repeated Prisoner's Dilemma tariff setting game. Uncertainty is introduced through country-specific shocks that influence the incentives to deviate from cooperation and that are only privately observable. To deal with such unanticipated shocks and to thus avoid breakdown of cooperation, the agreement can include a clause that allows a shock-affected country to escape from its liberalization commitments, without causing infinite reversion to the Nash outcome.

By introducing such an escape clause into the agreement, cooperation can be sustained for any degree of trade liberalization. In order to restrain the invocation of the escape clause, its use must be associated with some cost. In the model, a shock-affected country will therefore be allowed to apply its optimal tariff while incurring a pre-determined fixed cost for its temporary deviation. It is assumed that this cost is not sunk but transferred to the trading partner, thereby alleviating the damage of being exposed to the optimal tariff by the other country.

Under the optimal compensation cost scheme, the first-best compensation cost, or when that is not possible due to participation constraints, the maximum possible compensation cost given that constraint, is applied. It is shown that while the first-best cost for using the escape clause increases in the degree of liberalization, it will nevertheless be more frequently used when the degree of liberalization is increased. This is in accordance with the general perception that a higher degree of liberalization should be associated with a higher frequency of safeguards being invoked.

When the optimal compensation cost scheme is implemented, the expected per-period payoff increases for any given degree of liberalization, because the compensation cost does not only serve as a signal for a willingness to stick to the agreement, but also has an efficiency-enhancing effect. By attaching a cost to the use of the escape clause, the negative impact of its use on the trading partner is, at least partially, internalized. Thus, its use is restrained and, in case the first-best cost is applied, it will, in fact, only be used when its expected global effect is positive. In this case, the compensation cost provides the correct incentives for the use of the escape clause. Given the presence of hidden information, making the true loss stemming from being exposed to the escape clause unobservable to outsiders, the principle of reciprocity is thus shown to be optimal regarding how large compensation should be, since under the first-best solution, compensation equals the expected loss incurred by the escape clause. However, if the first-best cost cannot be implemented, the principle of reciprocity may not be upheld. In this case, compensation is lower than the expected loss incurred by the escape clause, and the escape clause will be used too frequently.

Besides sustaining the agreement and increasing expected per-period payoffs, there is a third positive effect of introducing the escape clause with compensation. When the optimal compensation cost scheme is applied, the scope for liberalization unambiguously increases for any given discount factor. Hence, freer trade can be obtained than in the absence of an escape clause.

The next section reviews the modeling of safeguards in the theoretical literature, and it briefly presents the various safeguard actions permitted under the legal framework of the General Agreement on Tariffs and Trade (GATT) and the World Trade Organization (WTO), with special emphasis on Article XIX of the GATT. Section 3 examines the case of the infinite repetition of a Prisoner's Dilemma game in the presence of unanticipated shocks. An escape clause is introduced in section 4, and its implications are investigated. The optimal design of an escape clause mechanism is derived for a special case in section 5. Section 6 concludes.

2 Background

2.1 Strategic Interaction in Trade Agreements

In the past two decades, a great variety of game theoretic models have been developed to explain the trade political setting within which countries operate. They can be divided into three broad categories:

– Cooperative games with some (implicitly assumed) enforcement mechanism.[1]

– Non-cooperative games without an enforcement mechanism.[2]

– Non-cooperative games with an (explicit) enforcement mechanism.[3]

There has been a continuing process starting out from static descriptions of trade policy in cooperative settings to increasingly sophisticated attempts at modelling enforcement in non-cooperative settings. During the same period, the Uruguay Round was completed and the WTO was founded, evolving from the GATT and including a reformed Dispute Settlement Procedure (DSP). Several models attempting at incorporating elements of the vast variety of different features of the WTO and its DSP (e.g. the Most-Favored-Nation (MFN) clause, reciprocity, concession diversion, and commensurate punishment) have recently appeared, e.g. Bagwell and Staiger (2002) and Ethier (2001).

This paper also starts out from the traditional game theoretic approach, modeling trade interaction between two countries as an infinitely repeated Prisoner's Dilemma tariff setting game. In the absence of cooperation, countries apply their respective optimal tariffs vis-à-vis each other and are hence stuck in a suboptimal Nash equilibrium. A Nash-superior cooperative outcome can, however, be achieved and sustained if there is sufficiently strong punishment against a deviator, e.g. infinite reversion to the Nash equilibrium. As long as there are no changes across periods, conditions for when cooperation is sustainable are thus easily derived. The present setting will, however, be modified such that each country is exposed to a shock influencing its incentives to deviate. By assuming these shocks to be non-observable to the trading partner, hidden information is introduced.

2.2 Models of Uncertainty and Escape

Strategic interaction in the presence of hidden information has mainly been addressed in the literature on collusion between firms, but also in the present context of trade policy. The impact of private information on strategic interaction between

[1] See, for example, Mayer (1981) and Riezman (1982).

[2] See, for example, Dixit (1987), Bagwell and Staiger (1990), and Riezman (1991).

[3] See, for example, Hungerford (1991), Kovenock and Thursby (1992), Maggi (1999), and Ludema (2001).

two countries was first addressed by Feenstra (1987) and Feenstra and Lewis (1991), which both apply cooperative one-stage settings. In Feenstra and Lewis (1991), the political pressure to restrict trade to which one government is exposed cannot be directly observed by its trading partner. In the presence of this informational asymmetry, the optimal outcome is to apply a tariff on imports exceeding some quota limit. Letting this limit vary, revenues and rents are allocated between the two countries in a way ensuring the truthful revelation of political pressures. Thus, transferring rents from trade restrictions can be regarded as having an informational role.[4]

The issue of cooperative outcomes in a repeated non-cooperative setting in the presence of hidden information has only recently been addressed in the literature on trade agreements.[5] Herzing (2005) investigates how the scope for liberalization in an infinitely repeated tariff setting game between two countries is affected by asymmetric information regarding the weight governments attach to present vis-à-vis future payoffs. Assuming away the possibility of any transfers across countries and any other cooperation-enhancing instruments, it is shown that self-destructive agreements, i.e. agreements prescribing a degree of liberalization associated with a strictly positive likelihood of cooperation breaking down, may be optimal ex ante.

Several contributions explicitly address the role of escape clauses in trade agreements. Ethier (2002) emphasizes that trade agreements are incomplete contracts negotiated when the future is uncertain. Ex post commitments under an agreement may become politically untenable, thereby necessitating the inclusion of a possibility to escape from these in order to sustain the agreement. A country exposed to the withdrawal of concessions under the escape clause by its trading partner will wish to be compensated by also withdrawing concessions. It turns out that the fundamental principle of reciprocity is optimal also with respect to the countermeasures permitted to be taken in response to the use of the escape clause. Hence, it is ex ante optimal to allow a country exposed to the escape clause by its trading partner

[4] Such transfers are similar both to the use of voluntary export restraints and the case with safeguards and monetary compensation.

[5] The impact of informational asymmetry on collusion between firms in an infinitely repeated non-cooperative setting has been analyzed by Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004).

to withdraw equivalent concessions.

A similar conclusion is drawn in Bagwell and Staiger (2002). They find that, in combination with the MFN clause, reciprocity as a governing principle is optimal not only in negotiations about liberalization, but also with respect to how ex post adjustments to the commitments should be conducted.[6] Central to their argument is that MFN and reciprocity together ensure that non-participants' terms of trade cannot be altered. When combined with the principle of reciprocity, the free-riding cost associated with MFN disappears. Thus, MFN and reciprocity provide a first-line of defense against the potential for opportunistic bilateral agreements, the second-line of defense being the possibility for governments of bringing non-violation nullification-of-impairment complaints.

Martin and Vergote (2004) model infinitely repeated strategic interaction between two countries, which are exposed to changes in political preferences that cannot be observed by the trading partner. In this setting, the antidumping instrument is used to let tariffs adjust to changes in these preferences. It is shown that in the absence of transfers or export subsidies, the former of which are rarely used and the latter of which are restricted under the GATT and the WTO, truthful revelation of political preferences can be achieved by letting present actions influence expected future payoffs through retaliation. The retaliatory use of antidumping may then improve welfare if static rules governing its use are adopted. Hence, when the use of some instruments are restricted, the strategic or retaliatory use of the remaining ones, such as antidumping, may be the most efficient way of dealing with hidden information.

Whether safeguards should be selectively applied is examined by Hochman (2004) under a different informational structure.[7] In a three-country-two-good setting, countries 1 and 2 export one good to country H, which is an exporter of the other good. In the presence of perfectly observable technology shocks independently affecting the production possibilities in countries 1 and 2 imports to country H will fluctuate, which may cause demand for protection. Two different safeguard regimes are compared. Under a selective safeguard, an importing country is permitted to

[6] See also Bagwell and Staiger (1999).

[7] The informational structure is similar to Rotemberg and Saloner (1986) and Bagwell and Staiger (1990).

apply the safeguard in a discriminatory fashion, while under a non-selective safeguard, its application must be on a MFN basis. Crucially, it is assumed that there exists a benefit from cooperating on non-economic objectives in countries H and 1. It is shown that a non-selective safeguard is optimal if this non-economic benefit is sufficiently large. If it is not large enough, tariffs on the efficient producer will be lower the larger is the trade-diversion effect. Compensation for using the safeguard is not necessary under non-selectivity, for it creates a cost for applying the safeguard. Moreover, expected global welfare is higher in the presence of a safeguard.

The model used here mainly draws on an approach used in Rosendorff and Milner (2001) who apply a two-stage game between two countries. In the first stage, in an international bargaining game, negotiations over the design of the institutional framework take place. In the second stage, there is an infinitely repeated trade policy game between countries, given the design of the institution. In each period, the political pressure for protection at home (and/or for more open markets abroad) is subject to a shock.[8] This shock can be seen as any exogenous and unanticipated change in the state of the world (unexpected price or supply shifts, changes in production technology, changes in a country's political institutions or preferences, changes in domestic political cleavages or alignments) that affects domestic firms' demand for, or ability to lobby for, protection of their markets. Furthermore, it is assumed that the two countries' shocks are stochastic and independently and identically distributed (i.i.d.), that in the current period each country knows its own state but not that of the other, and that both are equally uninformed about the weight values (at home and abroad) in all future periods.

In each period, the countries find themselves in a Prisoner's Dilemma. By deviating, a short-term gain can be made, but at the cost of infinite Nash reversion thereafter (given that grim-trigger strategies are applied). Ex ante, expectations are formed about the outcomes under mutual cooperation, deviation, being deviated against and the Nash equilibrium. There is assumed to be an upper bound to the short-term gain that can be made by deviating, and it is shown that for a sufficiently high discount factor, cooperation can be sustained forever. However, since coopera-

---

[8] In fact, a simplified version of the politically realistic objective function (PROF) proposed in Baldwin (1987) is applied.

tion may break down at lower discount factors, the possibility of exercising an escape clause at a fixed cost, allowing deviation for one period, is introduced. Having used the escape clause for one period, the country returns to the cooperative regime in the next period, having preserved its reputation as a cooperator. No supranational enforcement agency is necessary to make the escape-clause-using country pay the cost, since it wishes to preserve its credibility in the future. The institution merely serves as a verification agency, much as did the Law Merchants institution.[9]

It is shown that in the equilibrium with an escape clause, deviation can be avoided for any discount factor, as long as this cost is not larger than a threshold value increasing in the discount factor. Moreover, it is established that either there is an escape clause with a level of cost inducing enough cooperation and no breakdown such that the value of the game in an escape clause equilibrium is larger than that of the same game without an escape clause, or the cost of escape is too high and the escape clause equilibrium is the same as the grim-trigger equilibrium in the absence of an escape clause. The model predicts that greater domestic uncertainty, or situations where political leaders are more sensitive to unanticipated changes in political pressures, should be associated with a larger reliance on escape clauses. Hence, countries more sensitive to domestic pressures should be the main proponents and users of escape clauses, and certain issue-areas, such as exchange rate mechanisms and trade agreements, should be more likely to have escape clauses than others, due to their greater levels of uncertainty.

2.3 The Safeguard Provisions of the GATT

The term "escape clause" usually refers to the safeguard measures permitted under Article XIX of the GATT.[10] More generally, a safeguard is a provision allowing a WTO member to withdraw or cease to apply its normal obligations to protect certain overriding interests under specified conditions. Safeguard provisions, on the one hand, provide governments with the means to deviate from specific liberalization commitments in certain circumstances and thus have a safety-valve function and, on the other hand, facilitate the signing of a protection-reducing agreement and thus

[9] See Milgrom, North, and Weingast (1990).
[10] This subsection is based on World Trade Organization (1994a).

serve as an insurance mechanism (Hoekman and Kostecki, 1995).

It is possible to distinguish the various safeguard measures permitted under the GATT with regard to whether they are temporary or permanent and whether they are targeted at a specific industry or have economy-wide implications (Hoekman and Kostecki, 1995). The following table contains the safeguard provisions under the GATT, classified according to scope and duration.

	Temporary action	Permanent action
Industry-specific	Article VI Article XVIII: A and C Article XIX Article XXV	Article XXVIII
Economy-wide	Article XII Article XVIII: B	Article XX Article XXI

The various objectives of these safeguard provisions are summarized as follows in Hoekman and Kostecki (1995).

– Combating "unfair" trade by applying antidumping and countervailing duties (Article VI)

– Establishment of an industry, i.e. infant-industry protection (Article XVIII Sections A and C)

– Facilitating adjustment of an industry, i.e. emergency protection (Article XIX)

– Seeking a derogation (waiver) from specific GATT rules (Article XXV)

– Alleviating balance of payment problems (Articles XII and XVIII Section B)

– Allowing for renegotiation of tariff concessions (Article XXVIII)

– Achieving health, safety and related objectives (Article XX)

– Maintaining national security (Article XXI)

With regard to temporary and unexpected pressures stemming from sudden increases in imports, the GATT thus contains a range of measures for signatory countries that can be applied for these to be alleviated. A WTO member may temporarily restrict the imports of a product to protect a specific domestic industry from an increase in imports that is causing, or threatening to cause, serious injury to the industry. There are several articles in the GATT that can be invoked to that purpose. First, Article VI contains provisions for dealing with dumping and export subsidization that harm the industry of the importing country or that of another exporting country. Countries are allowed to impose antidumping duties to alleviate the damage of dumping, and countervailing duties are permitted to offset "any bounty and subsidy bestowed, directly, or indirectly, upon the manufacture, production or export of any merchandise" (Art VI §3). Second, Article XII allows a contracting party to restrict imports "in order to safeguard its external financial position and its balance of payments" (Art XII §1). Third, Article XVIII addresses the potential need for developing countries to take protective measures. Especially the promotion of "the establishment of particular industries with a view to raising the general standard of living of its people" (Art XVIII §3) is emphasized. Fourth, Article XIX provides the possibility of taking emergency actions on the imports of particular products. Finally, Article XXV addresses the possibility, "in exceptional circumstances not elsewhere provided for" in the agreement, for the signatory countries to "waive an obligation imposed upon a contracting party" (Art XXV §5) by the GATT.

In what follows, special attention will be given to Article XIX, since it specifically focuses on situations where a country suffers from sudden import surges seriously threatening domestic industries and thus, may be exposed to the temptation to break commitments made in the GATT. To prevent deviations from the agreement in such situations, Article XIX provides the possibility to temporarily suspend obligations. Article XIX §1(a) reads:

If, as a result of unforeseen developments and of the effect of the obligations incurred by a contracting party under this Agreement, including tariff concessions, any product is being imported into the territory of that contracting party in such increased quantities and under such conditions as to cause or threaten serious injury to domestic producers in that territory of like or directly competitive products, the contracting party shall be free, in respect of such product, and to the extent and for such time as may be necessary to prevent or remedy such injury, to suspend the obligation in whole or in part or to withdraw or modify the concession.

In Article XIX §2, the necessity of giving notice of any such action to the contracting parties as far in advance as possible is emphasized. Furthermore, the exporters of the products concerned shall be given the opportunity for consultation regarding the action. However, if delay of implementation of the action were harmful, provisional action without prior consultation would be allowed for, provided that consultation is effected immediately after taking such action. In case consultations do not end in agreement, the affected country is nevertheless permitted to take or continue the action, in which case "the affected contracting parties shall then be free, not later than ninety days after such action is taken, to suspend, upon the expiration of thirty days from the day on which written notice of such suspension is received by the contracting parties, the application to the trade of the contracting party taking such action, or ... to the trade of the contracting party requesting such action, of such substantially equivalent concessions or other obligations under this Agreement the suspension of which the contracting parties do not disapprove" (Art XIX §3:a). However, in case a country faces serious injury due to such action, it is "free to suspend, upon the taking of the action and throughout the period of consultation, such concessions or other obligations as may be necessary to prevent or remedy the injury" (Art XIX §3:b). Hence, there is a clear provision for retaliatory action against the application of a safeguard measure under Article XIX.

2.4 The Uruguay Round Agreement on Safeguards

The Tokyo Round (1973-1979) did not produce an agreement on a code of conduct governing the use of safeguard measures pursuant to Article XIX.[11] In the Uruguay Round (1986-1994), however, an agreement on safeguards was achieved.[12] With special regard to clarifying and reinforcing the disciplines of Article XIX, the need "to re-establish multilateral control over safeguards and eliminate measures that escape such control" is emphasized. The application of safeguard measures to a product is conditioned on that product being imported "in such increased quantities,

[11] This subsection is based on World Trade Organization (1994b).

[12] The Uruguay Round agreements consist of the Agreement Establishing the World Trade Organization and, annexed to it, the agreements on trade in goods (GATT), trade in services (GATS) and trade-related aspects of intellectual property rights (TRIPS), the dispute settlement understanding, the trade policy review mechanism and the plurilateral agreements, as well as the schedules of commitments. The multilateral agreements on trade in goods (annex 1A) contain, among others, the Agreement on Safeguards.

absolute or relative to domestic production, and under such conditions as to cause or threaten to cause serious injury to the domestic industry that produces like or directly competitive products" (Art 2). "Serious injury" is defined as "significant overall impairment in the position of a domestic industry" (Art 4:1a), while "threat of serious injury" is defined as clearly imminent serious injury, the determination of which "shall be based on facts and not merely on allegation, conjecture or remote possibility" (Art 4:1b). When determining whether increased imports are causing or threatening to cause serious injury, particular emphasis should be put on "the rate and amount of the increase in imports of the product concerned in absolute and relative terms, the share of the domestic market taken by increased imports, changes in the level of sales, production, productivity, capacity utilization, profits and losses, and employment" (Art 4:2a).[13] Safeguard measures should be applied "only to the extent necessary to prevent or remedy serious injury and to facilitate adjustment" (Art 5:1) and be on a MFN basis, although selective applications are permitted under Article 5:2b.[14]

Safeguard measures are only permitted for a period not exceeding four years (Art 7:1), except under special circumstances (Art 7:2); the total period of application shall not exceed eight years, however (Art 7:3). Whenever a safeguard measure can be expected to be applied for more than one year, "the Member applying the measure shall progressively liberalize it at regular intervals during the period of application" (Art 7:4). Moreover, to prevent circumvention of the requirement of liberalization of safeguard measures, it is not permitted to apply a safeguard measure "for a period of time equal to that during which such measure had been previously applied, provided that the period of non-application is at least two years" (Art 7:5).[15] These clearly defined time limits on the use of a safeguard measure stand in sharp contrast to the previous, somewhat vague constraint for the application of a measure to be

[13] Schott (1994) notes that the provisions for serious injury establish a much higher threshold than the "material injury" standard for antidumping and countervailing duties outlined in Article VI.

[14] According to Schott (1994), these selective applications seem to have been aimed at exporters in Eastern Europe and Asia, in particular China.

[15] An exception is made if the safeguard measure has a duration of 180 days or less, at least one year has elapsed since the introduction of the initial measure and such a measure has not been applied more than twice during the five years preceding the date of introduction of the measure (Art 7:6).

"temporary".

In order to alleviate the effects of exercising a safeguard measure, it is required that "a substantially equivalent level of concessions and other obligations to that existing under GATT 1994" is maintained between the measure-taking party and the affected parties, which may be achieved through "any adequate means of trade compensation for the adverse effects of the measure on their trade" (Art 8:1). In case agreement is not reached within 30 days, an affected exporting party has the right to suspend the application of "substantially equivalent concessions or other obligations under GATT 1994" (Art 8:2), but not "for the first three years that a safeguard measure is in effect, provided that the safeguard measure has been taken as a result of an absolute increase in imports and that such a measure conforms to the provisions of this Agreement" (Art 8:3).

Developing countries are granted special treatment. Safeguards are not allowed against a developing country "as long as its share of imports of the product concerned in the importing Member does not exceed 3 per cent, provided that developing country Members with less than 3 per cent import share collectively account for not more than 9 per cent of total imports of the product concerned" (Art 9:1). Moreover, a developing country has the possibility of extending a safeguard measure up to ten years, inclusive of extensions, and can reintroduce a safeguard measure against a product after a period of time equal to only half that during which the previous was applied, given that two years have passed (Art 9:2).

To stem the use of non-orderly quantitative restrictions, "voluntary export restraints, orderly marketing arrangements or any other similar measures on the export or the import side" are explicitly prohibited (Art 11:1b), which has been considered as one of the greatest achievements of the Uruguay Round.[16]

Schott (1994) notes that before the conclusion of the Uruguay Round, Article XIX measures had to be non-discriminatory and affected exporters had the right to claim compensation or seek authorization for retaliation. Governments thus often preferred other less costly and more flexible safeguard measures, either because they sought to exempt certain countries, or to avoid the need for compensation.[17]

[16] See, for example, Schott (1994) and Hoekman and Kostecki (1995).

[17] Feenstra (1987) emphasizes that limitations on the use of temporary import restrictions under

When assessing the results of the Uruguay Round, Schott (1994) emphasizes the ban on VERs as one of the major achievements, but criticizes the weak incentives for adjustment resulting from the generous durations of safeguard actions, the failure to remove the justification for balance of payments safeguard measures imposed by developing countries and the failure to discipline their use. He notes that the removal of the threat of retaliation for three years, the ban on alternative "gray area" measures such as VERs, the omission of an adjustment requirement for at least four years and the possibility for a selective application of safeguards seem to have been aimed at encouraging the use of Article XIX. However, the reduced risk of retaliation is balanced by the relatively rigid serious injury requirement, the MFN requirement in most cases and the constraint on actions vis-à-vis developing countries. Therefore, it is likely that safeguard measures will continue to play a minor role, at least as compared to antidumping actions.

2.5 Economic Implications of Safeguard Measures

Hoekman and Kostecki (1995) note that safeguard actions distribute income from consumers to import-competing and/or foreign exporting industries. Whatever the political rationale for safeguard instruments, their mere existence may reduce competitive pressure on domestic import-competing firms, e.g. by raising prices or reducing the incentives to innovate. Scope may also exist for the capture and abuse of such procedures by import-competing interests further strengthening such effects. Hence, the gains from the liberalization negotiated under multilateral trade negotiations or unilaterally implemented are reduced for certain sectors or for the economy as a whole. In so far as the cause of an import-competing industry's problems lies in a shift in comparative advantage, protection is generally an inappropriate policy for creating adjustment. Moreover, safeguards are usually economically inefficient, since the costs for consumers are almost always larger than the benefits accruing to the protected industry. Furthermore, industries can be expected to exploit substitution possibilities across instruments if these exist, thereby making it more difficult

---

Article XIX of the GATT may be ineffective when one country cannot actually verify the conditions faced by an industry in the other country. Hence, incentives to misrepresent the conditions faced by domestic industries to obtain protection may arise in the presence of incomplete information.

for governments to control trade policy.

3 A Prisoner's Dilemma Tariff Setting Game

3.1 The Game under Certainty

The setting is as follows. There are two countries, each exporting one good to the other, but perfectly symmetric in all other respects. Each country's payoff W is a function of its own tariff t and the foreign tariff t^*, i.e. $W = W(t, t^*)$. There exists a best-response function $t_{BR}(t^*) = \arg\max_t W(t, t^*)$. W increases in t for $t < t_{BR}(t^*)$, while it decreases in t for $t > t_{BR}(t^*)$ (as long as trade takes place). Hence, for any given t^*, there is a unique t maximizing W (as long as trade takes place). W is monotonously falling in t^* (as long as trade takes place).

A simple and frequently used way of modeling trade policy is to regard tariff setting interaction between two countries as an infinite repetition of a Prisoner's Dilemma game, where both countries can choose between cooperating or deviating. In fact, there are two stages. In the first stage, before the infinitely repeated game begins, the two countries choose a cooperative tariff level t_C from a continuum and agree on how deviations should be punished. In the second stage, the infinitely repeated game is played. Once it starts, each country will choose between implementing the agreed-upon cooperative tariff and applying the optimal tariff $t_D = t_{BR}(t_C)$ vis-à-vis the other country. Actually, a country can choose its tariff level from a continuum. However, assuming that setting the tariff different from t_C is regarded as a deviation, a country's choice will, in fact, be binary, i.e. between applying t_C and t_D. Under perfect symmetry, the per-period payoff under cooperation is given by $W_C = W(t_C, t_C)$. A country deciding to break its commitment by applying the optimal tariff vis-à-vis its trading partner obtains the payoff $W_D = W(t_D, t_C)$, while its trading partner receives the sucker's payoff $W_S = W(t_C, t_D)$. In the absence of cooperation, both countries apply the optimal tariff vis-à-vis each other, i.e. the Nash tariff rate $t_N = t_{BR}(t_{BR})$, and both receive the payoff $W_N = W(t_N, t_N)$. The chosen cooperative tariff level t_C does not only directly define payoff under cooperation (W_C), but also indirectly, via the best-response function, defines payoffs of deviation (W_D) and being deviated-against (W_S).

There exists a unique tariff level $t_C^{opt} < t_N$ that maximizes $W_C = W(t_C, t_C)$. W_C increases at a decreasing rate in t_C for $t_C < t_C^{opt}$, while it decreases at an increasing rate for $t_C > t_C^{opt}$. It immediately follows that there exists a $t' < t_C^{opt}$, such that $W_C = W_N$ for $t_C = t'$. Thus, $W_C > W_N$ if and only if $t' < t_C < t_N$.

W_S increases at a decreasing rate in t_C for $t_C < t_N$ and decreases at an increasing rate in t_C for $t_C > t_N$, while W_D decreases in t_C. Letting t_C decrease below t_N thus leads to a monotonous increase in W_D. The increase in W_D is equal to that in W_C (i.e. W_D and W_C are tangent) at $t_C = t_N$ and unambiguously stronger as t_C falls below t_N. Thus, $W_D - W_C$ increases, and does so at an increasing rate as t_C decreases.

Define

$$\tau \equiv \frac{W_D - W_C}{W_D - W_N}.$$

It is easily shown that $\lim_{t_C \to t_N} \tau = 0$ and that τ increases as t_C decreases below t_N. Hence, τ can be seen as a measure of trade liberalization. A low value of t_C corresponds to a high value of τ and thus, a high degree of trade liberalization. Since $W_C^{opt} > W_N$, the optimal degree of liberalization τ^{opt} is strictly smaller than unity and, because $W_C = W_N$ for $t_C = t'$, the degree of liberalization τ' corresponding to $t_C = t'$ equals unity.

The relevant range of cooperative tariffs to be considered are those yielding Nash-superior cooperative payoffs, i.e. $t' < t_C < t_N$. In this range, it is the case that $W_D > W_C > W_N > W_S$. (For $t_C = t_N$, it is obviously the case that $W_D = W_C = W_S = W_N$) The per-period payoff matrix for the symmetric case, given below, is thus of Prisoner's Dilemma type.

	Cooperate	Deviate
Cooperate	W_C, W_C	W_S, W_D
Deviate	W_D, W_S	W_N, W_N.

A cooperative outcome in every period can be maintained if countries stick to a grim-trigger strategy, i.e. any deviation by one country will be punished by infinite reversion to the Nash equilibrium. Cooperation is sustainable, if and only if the cost of deviating outweighs the one-period gain from deviating, i.e. if and only if

$$W_D - W_C \leq \frac{\beta}{1-\beta}(W_C - W_N), \tag{1}$$

where β is the discount factor. The left-hand side represents the short-term (one-period) gain from deviation, while the right-hand side represents the expected long-term loss from deviation. Assuming $W_C > W_N$, rearranging terms yields the following relationship between the degree of liberalization and the discount factor

$$(1) \Leftrightarrow \frac{W_D - W_C}{W_C - W_N} \leq \frac{\beta}{1-\beta} \Leftrightarrow \frac{\tau}{1-\tau} \leq \frac{\beta}{1-\beta} \Leftrightarrow \tau \leq \beta. \tag{1'}$$

This equation tells us that to sustain cooperation, t_C can only be reduced to the degree that τ does not exceed β. A higher discount factor implies that the upper bound for liberalization increases, and it is thus possible to sustain a lower t_C. The restriction given by (1') is incorporated into the negotiations concerning the cooperative tariff level and thus imposes an upper limit on liberalization. By introspection of (1') it immediately follows that it is always possible to find some $t_C < t_N$ sustainable for $\beta > 0$, and that the optimal cooperative tariff can be sustained if the discount factor is sufficiently large.

3.2 Introducing Uncertainty and Asymmetric Information

So far, it has been assumed that there is perfect certainty regarding payoffs both in the present and all future periods. If there is certainty about the payoffs in each period, infinite cooperation can easily be sustained, as long as (1) holds. However, unanticipated events may occur. In what follows, it will be assumed that in every period, each country is affected by the exogenous shocks ε and ε^*, respectively. These shocks are assumed to be identically and independently distributed across periods and between countries and not observable to the trading partner. Hence, each country has perfect knowledge about the shock to which it is exposed, but it knows nothing about the shock its trading partner is presently experiencing, neither does it know anything about the shocks that will occur both at home and abroad in future periods. Countries must thus take their decisions in the presence of hidden information.

The shocks are assumed to influence the demand for protection. The underlying reason for increased demand for protection could be surges in import volumes, like in Bagwell and Staiger (1990), with the additional assumption of non-observability by the trading partner. Another possibility could be shifts in the political clout of the import-competing sector, like in Rosendorf and Milner (2001) which applies a simplified version of Baldwin's (1987) PROF. A third possibility could be the government becoming more impatient and thus, attributing more weight to present vis-à-vis future payoffs, like in Herzing (2005).[18] However, the present analysis focuses on how randomly determined incentives to deviate affect cooperation, and not on the exact nature of the underlying shock creating fluctuations in these incentives.

There are three channels through which the different payoffs of the game are affected by the stochastic variables. First, the payoff functions are affected such that $W^{\varepsilon} = W(t, t^*; \varepsilon)$. Second, the own best-response function is affected such that $t_D^{\varepsilon} = t_{BR}^{\varepsilon}(t_C)$. And third, the trading partner's best-response function is affected such that $t_D^{\varepsilon^*} = t_{BR}^{*\varepsilon^*}(t_C)$. Thus, while the payoff under cooperation will only be affected through the first channel, i.e. $W_C^{\varepsilon} = W(t_C, t_C; \varepsilon)$, the payoff under deviation will also be affected through the second channel, i.e. $W_D^{\varepsilon} = W(t_D^{\varepsilon}, t_C; \varepsilon)$. The sucker's payoff and the Nash payoff will both be affected through the third channel. Whereas the sucker's payoff is not affected through the second channel, i.e. $W_{S,\varepsilon^*}^{\varepsilon} = W(t_C, t_D^{*\varepsilon^*}; \varepsilon)$, the Nash payoff is affected through all three channels, i.e. $W_{N,\varepsilon^*}^{\varepsilon} = W(t_D^{\varepsilon}, t_D^{*\varepsilon}; \varepsilon)$.

For tractability, it will be assumed that the best-response functions are unaffected by these exogenous shocks

$$\text{(A1)} \qquad t_{BR}^{\varepsilon} = t_{BR} \qquad \text{for any } \varepsilon.$$

By making this simplifying assumption, the analysis is solely focused on the effects of a shock through the first channel. Thus, the shock experienced by the trading partner has no direct effect on the own payoff. It is only the domestic shock that affects payoffs, i.e. $W^{\varepsilon} = W(t, t^*; \varepsilon)$ and $W^{*\varepsilon^*} = W(t, t^*; \varepsilon^*)$. However, the shock of the trading partner has an effect on its strategic decision on whether to

[18] In fact, in the present context of hidden information, such uncertainty is similar to introducing uncertainty in Baldwin's (1987) PROF, as demonstrated in Herzing (2005).

continue cooperation.

Next, it will be assumed that the difference in payoffs between deviating and cooperating is independent of the action taken by the trading partner

$$\text{(A2)} \qquad W(t_D, t^*; \varepsilon) - W(t_C, t^*; \varepsilon) = \text{constant for all } t^*.$$

Thus, the one-period gain of deviating is independent of the trading partner's action. This implies in particular that $W_N^\varepsilon - W_S^\varepsilon = W_D^\varepsilon - W_C^\varepsilon$. Making this assumption simplifies matters since there is no need to consider what action will be taken by the trading partner. Therefore, any decision will solely depend on the domestic shock.

Both (A1) and (A2) can be justified in the context of a partial equilibrium framework, where payoffs can be additively separated such that terms solely depending on t and t^*, respectively, are obtained.

To sustain cooperation, the expected outcome under deviation must be smaller than the expected outcome under cooperation. Let V_C be the continuation value if cooperation is sustained in the present period, and let V_D be the continuation value if deviation occurs. Let p be the probability that the trading partner chooses to cooperate (p can be seen as a prior regarding the trading partner's choice)

$$\begin{aligned} C \succcurlyeq D \Leftrightarrow\ & p[W_C^\varepsilon + \beta V_C] + (1-p)[W_S^\varepsilon + \beta V_D] \\ & \geq p[W_D^\varepsilon + \beta V_D] + (1-p)[W_N^\varepsilon + \beta V_D] \\ \Leftrightarrow\ & \beta p[V_C - V_D] \geq p[W_D^\varepsilon - W_C^\varepsilon] + (1-p)[W_N^\varepsilon - W_S^\varepsilon] \\ \overset{(A2)}{\Leftrightarrow}\ & W_D^\varepsilon - W_C^\varepsilon \leq \beta p[V_C - V_D]. \end{aligned} \tag{2}$$

In the absence of shocks, (2) is equivalent to equation (1) with an implied probability of cooperating equal to unity ($p = 1$). If shocks occur, however, the decision to cooperate depends on the size of the one-period gain to be made by deviating. Since present shocks are assumed not to have any impact on expectations for future payoffs, the continuation values are based on expected values independent of the present situation. Hence, the right-hand side of (2) is constant, and it is the size of the left-hand side that determines the choice of whether to cooperate. What is of importance is therefore the difference in payoffs between deviating and cooperating under the shock.

Next, it will be assumed that the exogenous variable affects the incentive to deviate in the following way

$$\text{(A3)} \qquad W_D^\varepsilon - W_C^\varepsilon = \varepsilon[W_D - W_C],$$

where $\varepsilon \in [0, \infty]$ is distributed according to the commonly known density function φ, which is assumed to be strictly positive for $\varepsilon \in (0, \infty)$ and stable over time.[19] The associated cumulative distribution function is denoted by Φ, and the expected value of ε is assumed to equal one.[20] The expected value of the one-period gain from deviating is thus equal to that in the game without shocks.

By making this assumption regarding the effect of the exogenous variable ε, it is taken into account that what is of importance is not absolute realizations of W_C^ε and W_D^ε, but how the one-period gain from deviating is affected. Thus, a country getting a lower payoff than expected under cooperation may nevertheless not opt for deviating, if the expected payoff under deviation is even lower. And, by the same token, a country getting a higher payoff from cooperation than expected may nonetheless be inclined to deviate, if the gain from deviating is even higher. Using (A2) and (A3), it immediately follows that

$$W_N^\varepsilon - W_S^\varepsilon = W_D^\varepsilon - W_C^\varepsilon = \varepsilon[W_D - W_C] = \varepsilon[W_N - W_S].$$

By defining the distribution function over the interval $[0, \infty]$, no arbitrary assumption is made regarding the upper bound for the size of the one-period gain from deviating. Thus, the one-period gain from deviating can potentially take on extremely large values, albeit with infinitesimally small probabilities. By letting the shock enter the equation multiplicatively, account is taken of the impact of a shock being different for different cooperative tariff levels. Since the one-period gain from deviating is higher the lower the cooperative tariff is set, the same shock will have a stronger impact at a lower cooperative tariff level. If the cooperative tariff is set

[19] One possible distribution could be a log-linear distribution function.

[20] While assumption (A3) is made for analytical convenience, it is not unfounded. In fact, it is easy to construct a correspondence to Herzing (2005), where the weight attributed to present payoffs is assumed to be stochastically determined.

equal to the Nash tariff ($t_C = t_N$), the payoffs under cooperation, deviation and being deviated against are all equal to the Nash payoff ($W_C = W_D = W_S = W_N$). In this case, the one-period gain from deviating will naturally be zero, independent of the size of the shock ε, which is captured by (A3). Furthermore, for any cooperative tariff level below the Nash tariff, a sufficiently large shock will make deviation worthwhile. Thus, in this setting, cooperation will not be sustainable for any $t_C < t_N$ if an unanticipated, sufficiently large shock occurs.

For $\tau = 0$, both sides of (2) equal zero. Hence, the country will be indifferent between cooperation and deviation. When $\tau > 0$, the left-hand side of (2) is strictly positive. Plugging (A3) into (2), the following condition for when cooperation will be preferred to deviation is obtained

$$C \succcurlyeq D \Leftrightarrow \varepsilon \leq \beta p \frac{V_C - V_D}{W_D - W_C} \equiv \eta.$$

Cooperation will thus be chosen as long as the shock is smaller than a threshold value that depends on the discount factor, the prior, the degree of liberalization and the implied continuation values. However, if a sufficiently large shock, exceeding this threshold value, occurs, deviation will be chosen and cooperation will break down. The continuation values are as follows (see the Appendix for derivation)

$$V_D = \frac{W_N}{1-\beta} \tag{3}$$

$$V_C = V_D + \frac{1}{1-\beta p^2}[p(W_D - W_N) - (W_D - W_C)\int_0^{\eta} \varepsilon d\Phi]. \tag{4}$$

By plugging the continuation values into the above equation, the following equation for η is obtained

$$\eta = \frac{\beta p}{1-\beta p^2}[\frac{p}{\tau} - \int_0^{\eta} \varepsilon d\Phi]. \tag{5}$$

Since η enters both the left-hand side (LHS) and the right-hand side (RHS) of (5), it is not straightforward whether and how (5) can be solved.

3.3 Consistent Solutions under Symmetry

A country will not know what action will be taken by its trading partner, but it can estimate the likelihood of cooperation, which must be consistent in the sense of being equal to the implied probability of choosing cooperation. By symmetry, which implies that the probability of choosing cooperation is the same for both countries, this probability is therefore given by $p = prob(\varepsilon \leq \eta) = \Phi(\eta)$. Plugging this into equation (5) gives us the following condition for consistent solutions

$$\eta = \frac{\beta\Phi(\eta)}{1-\beta\Phi^2(\eta)}[\frac{\Phi(\eta)}{\tau} - \int_0^\eta \varepsilon d\Phi] \equiv f(\eta). \tag{6}$$

The left-hand side of (6) takes on values in the interval $[0, \infty]$, while the right-hand side takes on values in a positive, bounded interval.

Lemma 1 *There exists at least one solution ($\eta = 0$) to equation (6).*

Proof. Since $f(0) = 0$, it immediately follows that $\eta = 0$ is a solution to (6). ■

Letting η go to infinity, f converges to $\frac{\beta}{1-\beta}\frac{1-\tau}{\tau}$. There may or may not exist further solutions, depending on the discount factor and the degree of liberalization.

By increasing the discount factor, $f(\eta)$ increases for any given η. Since for $\Phi(\eta)$ close to unity letting β approach unity will make the term $\frac{\beta\Phi(\eta)}{1-\beta\Phi^2(\eta)}$ very large, it is the case that for a sufficiently high discount factor, there will be values of η, for which $f(\eta) \geq \eta$. By decreasing the discount factor, $f(\eta)$ decreases for any given η, and eventually $f(\eta) < \eta$ for all $\eta > 0$. Thus, there exists a threshold value of β, below which there exist no further solutions to equation (6).

The reasoning is similar for the chosen degree of liberalization. Letting τ go to zero, $f(\eta)$ becomes infinitely large for $\eta > 0$. As τ increases, $f(\eta)$ decreases for any given $\eta > 0$. Eventually, when τ is sufficiently large, $f(\eta) < \eta$ for all $\eta > 0$. In this case, there will exist no further solution to equation (6). However, for τ sufficiently close to zero, there will at least be an interval of values for η, for which $f(\eta) \geq \eta$.

Define η' as the largest solution to (6), i.e. $\eta' \equiv \max\{\eta | \eta = f(\eta)\}$. The following lemma demonstrates why it is reasonable to regard η' as the preferred threshold value.

Lemma 2 *η' is renders the highest continuation value among all solutions to (6).*

Proof. See the Appendix. ■

It is reasonable to assume that both countries share an interest in fostering a belief that η', being the solution rendering the highest expected payoff, is the threshold value applied. Hence, η' will be referred to as the relevant solution to (6).

By assumption (A3) it is the case that for any strictly positive degree of liberalization, and given that $\beta < 1$, cooperation cannot be sustained forever. As soon as a sufficiently strong shock occurs, deviation will be chosen. Thus, cooperation will break down in finite time.[21] It is straightforward that under (A3), an implied likelihood of cooperation being chosen equal to unity is only possible if $\beta = 1$ or $\tau = 0$.

It is important to emphasize the significance of asymmetric information in this setting. If shocks were perfectly observable, a result similar to that of Bagwell and Staiger (1990) would be obtained. In the absence of the possibility to misrepresent the true state of a country, an agreement could prescribe temporary (i.e. one period lasting) Nash reversion for sufficiently high realizations of ε (and/or ε^*).

Alternatively, unilateral deviation could be permitted for sufficiently high realizations of ε. The deviated-against country could then be compensated by being permitted to deviate in some future period. Such a solution would be similar to the collusive schemes obtained in Athey and Bagwell (2001), under which compensation for a current-period advantage by one firm is in the form of a future market-share concession instead of a side payment by this firm.[22] However, such intertemporal compensation schemes will not be considered in the present analysis.[23]

To conclude, while it is not possible to infinitely sustain cooperation given a certain discount factor, it can be sustained for sufficiently low degrees of liberalization,

[21] While this result crucially hinges on the fact that infinitely large realizations of ε are allowed for, it could nevertheless be true that the ex ante likelihood of cooperation is smaller than unity for sufficiently high degrees of liberalization, even in the presence of an upper bound on ε (see Herzing (2005)).

[22] Indeed, future market-share favors are strictly preferred to side payments, unless the latter are perfectly efficient (i.e. detection is of no concern), which is the case here.

[23] The collusive schemes in Athey and Bagwell (2001) are actually derived under hidden information, suggesting that such a solution could be feasible also in the present setting.

as long as the realizations of the stochastic variable are smaller than a threshold value. Cooperation will, however, break down as soon as this threshold value is exceeded. When the degree of liberalization is too large, cooperation will break down instantly for in this case, each country will anticipate that its trading partner will deviate and thus also decide to deviate.

4 Introducing an Escape Clause

If both countries know in advance that they might be exposed to the temptation to deviate from cooperation, they might ex ante agree on a mechanism for dealing with situations where shocks inducing changes in the payoff occur.

In what follows, a modification of the approach taken by Rosendorff and Milner (2001) will be applied. Using the setting described above, an agreement on cooperation could then include the possibility of making use of an escape clause, in case a country faces a severe shock, thereby allowing this country to offset the effects of the shock by temporarily deviating, i.e. during the period during which the shock occurs, and then return to the cooperative regime.

However, in the presence of asymmetric information it is impossible for any outsider to determine whether the use of the escape clause is legitimate. Hence, it is necessary that the country using the escape clause and thus temporarily deviates incurs some cost, either by imposing the cost upon itself or by being exposed to some action by its trading partner. What is of importance is that by exercising the escape clause and thus incurring a cost, a shock-affected country signals its willingness to return to the cooperative regime by making a voluntary concession. In the absence of any such cost, it becomes rational for each country to always apply the escape clause. While the agreement would be preserved under such a scenario, the outcome would be the same as in the absence of any agreement. Hence, to ensure that cooperation is chosen at least in some periods, the use of the escape clause must be associated with some cost.

Hence, instead of deviating, receiving the one-period payoff W_D and thereby making further cooperation impossible, the country deviates and incurs a fixed cost

$F(\varepsilon) > 0$, thereby receiving the payoff $W_D - F(\varepsilon)$.[24] The cost incurred by the country exercising the escape clause is assumed to translate into a benefit of exactly the same size for its trading partner.[25] Therefore, the escape-clause-exposed country receives the sucker's payoff W_S, plus compensation $F(\varepsilon)$.[26]

The assumption of instantly incurring a compensatory cost is made for analytical tractability. Actually, as emphasized in subsection 2.2, a country is permitted to apply Article XIX of the GATT without initially incurring a cost; only when it continues to apply the escape clause will it be obliged to compensate its trading partner. How the present model can be modified to more realistically reflect the legal possibilities of invoking the escape clause is addressed in the concluding section.

The payoff matrix now looks as follows.[27]

	C^*	EC^*	D^*
C	W_C^ε, $W_C^{\varepsilon^*}$	$W_S^\varepsilon + F(\varepsilon^*)$, $W_D^{\varepsilon^*} - F(\varepsilon^*)$	W_S^ε, $W_D^{\varepsilon^*}$
EC	$W_D^\varepsilon - F(\varepsilon)$, $W_S^{\varepsilon^*} + F(\varepsilon)$	$W_N^\varepsilon + F(\varepsilon^*) - F(\varepsilon)$, $W_N^{\varepsilon^*} + F(\varepsilon) - F(\varepsilon^*)$	$W_N^\varepsilon - F(\varepsilon)$, $W_N^{\varepsilon^*} + F(\varepsilon)$
D	W_D^ε, $W_S^{\varepsilon^*}$	$W_N^\varepsilon + F(\varepsilon^*)$, $W_N^{\varepsilon^*} - F(\varepsilon^*)$	W_N^ε, $W_N^{\varepsilon^*}$

The sequence of events during a period is as follows:

1. At the beginning of a period, both countries experience independent shocks that are unobservable to the trading partner.

2. Both countries determine, independently, what policy they will apply, the options being cooperation, use of the escape clause, and deviation.

[24] Note that for simplicity, the cost for deviating is assumed to be independent of the degree of deviation. Hence, optimal deviation will always be chosen.

[25] Such an assumption is strong, especially if compensation were in the form of a decrease in some other tariff or an increase in the other country's tariff. It is not unrealistic, however. It is valid if compensation is purely monetary, i.e. the country exercising the escape clause is required to transfer money to its trading partner, a case that will be more closely examined in the next section.

[26] Note that, unlike Rosendorff and Milner (2001), this cost is not sunk; the trading partner receives no compensation in their model.

[27] Note that for $F(\varepsilon) > 0$, the only Nash equilibrium of the one-shot game is (D, D*).

3. Both countries implement their policies, and the period begins.
4. At the end of the period, the implemented policies are verified. Any deviation by one country is regarded as a breach of the agreement and will therefore lead to a breakdown of the cooperative regime.

First, the threshold value, below which cooperation is preferred to deviation, is derived. Let p_C and p_{EC} be the probabilities that the trading partner opts for cooperation and exercising the escape clause, respectively, and let V_C^{EC} be the continuation value of the game if no deviation occurs in the present period. As before, $V_D = \frac{W_N}{1-\beta}$ is the continuation value if deviation occurs and it is followed by infinite Nash reversion

$$\begin{aligned}
& C \succcurlyeq D \\
& \quad \Leftrightarrow p_C[W_C^\varepsilon + \beta V_C^{EC}] + p_{EC}[W_S^\varepsilon + \beta V_C^{EC}] + (1 - p_C - p_{EC})[W_S^\varepsilon + \beta V_D] \\
& \qquad \geq p_C[W_D^\varepsilon + \beta V_D] + p_{EC}[W_N^\varepsilon + \beta V_D] + (1 - p_C - p_{EC})[W_N^\varepsilon + \beta V_D] \\
& \quad \Leftrightarrow p_C[W_C^\varepsilon - W_D^\varepsilon] + (1 - p_C)[W_S^\varepsilon - W_N^\varepsilon] + \beta(p_C + p_{EC})[V_C^{EC} - V_D] \geq 0 \\
& \quad \Leftrightarrow W_C^\varepsilon - W_D^\varepsilon + \beta(p_C + p_{EC})[V_C^{EC} - V_D] \geq 0 \\
& \quad \Leftrightarrow \varepsilon[W_C - W_D] + \beta(p_C + p_{EC})[V_C^{EC} - V_D] \geq 0 \\
& \quad \Leftrightarrow \varepsilon \leq \beta(p_C + p_{EC})\frac{V_C^{EC} - V_D}{W_D - W_C}.
\end{aligned}$$

Next, a condition for the compensation cost ensuring that deviation renders a lower outcome than the use of the escape clause is derived

$$\begin{aligned}
& EC \succcurlyeq D \\
& \quad \Leftrightarrow p_C\{W_D^\varepsilon - F(\varepsilon) + \beta V_C^{EC}\} + p_{EC}\{W_N^\varepsilon - F(\varepsilon) + F(\varepsilon^*) + \beta V_C^{EC}\} \\
& \qquad\qquad + (1 - p_C - p_{EC})\{W_N^\varepsilon - F(\varepsilon) + \beta V_C^{EC}\} \\
& \qquad \geq p_C\{W_D^\varepsilon + \beta V_D\} + p_{EC}\{W_N^\varepsilon + F(\varepsilon^*) + \beta V_D\} \\
& \qquad\qquad + (1 - p_C - p_{EC})\{W_N^\varepsilon + \beta V_D\} \\
& \quad \Leftrightarrow F(\varepsilon) \leq (p_C + p_{EC})\beta[V_C^{EC} - V_D].
\end{aligned}$$

If this condition is satisfied for $\varepsilon > \beta(p_C + p_{EC})\frac{V_C^{EC} - V_D}{W_D - W_C}$, deviation will never be an option, and hence $p_{EC} = 1 - p_C$. In this case, the above conditions can be expressed as follows

$$C \succcurlyeq D \Leftrightarrow \varepsilon \leq \beta \frac{V_C^{EC} - V_D}{W_D - W_C} \equiv \overline{\eta} \tag{7}$$

$$EC \succcurlyeq D \Leftrightarrow F(\varepsilon) \leq \beta[V_C^{EC} - V_D] = \overline{\eta}[W_D - W_C] \equiv \overline{F}. \tag{8}$$

Condition (8) can be seen as a participation constraint. By imposing an upper limit on the cost for using the escape clause, it is ensured that deviation from the agreement can never be a preferred option. Hence, to avoid deviation, the cost for using the escape clause cannot be too high for shocks above the threshold value $\overline{\eta}$ given by (7). More specifically, this cost cannot be larger than the discounted future gain from sustaining cooperation vis-à-vis infinite Nash reversion.

In conclusion, to avoid deviations in the presence of unanticipated, temporary shocks, an agreement could include an escape clause, allowing a shock-affected country to temporarily deviate while incurring a cost. The escape clause should be used whenever the one-period gain from deviating exceeds the discounted future gain from sustaining the agreement vis-à-vis infinite Nash reversion.

5 The Optimal Design of an Agreement with an Escape Clause and Monetary Compensation

5.1 The Incentive Problem

Implementing an escape clause cost scheme $F(\varepsilon)$, exactly prescribing the cost associated with exercising an escape clause at a certain shock level, will depend on the observability of shocks. If both ε and the implementation of $F(\varepsilon)$ are perfectly observable, any transfer scheme $F(\varepsilon)$ satisfying (8) for $\varepsilon > \overline{\eta}$ will ensure that deviation never occurs. Here, information about the shock is assumed to be asymmetric, however,which leads to incentive problems. Depending on what compensation mechanism is chosen, there may be strong incentives to under- or overestimate the size of a shock, both for the shock-affected country and the escape-clause-exposed country.

There is a great degree of freedom in deciding the shape of the compensation scheme $F(\varepsilon)$, provided that the above conditions are met. As can be expected, the range for ε where cooperation will be chosen increases in F, i.e. the cost associated with exercising the escape clause.[28] A country will prefer F to be low to be able to increase the range within which the escape clause can be exercised and not to have to pay so much compensation in case it faces strong incentives to deviate. However, in case its trading partner faces a shock, it will prefer F to be high to decrease the range within which the escape clause can be exercised and get a high compensation. Ex ante, when a country knows it can be in either position in the future, these two effects on the expected payoff must be weighed against each other.[29]

If compensation is independent of the size of a shock, i.e. $F(\varepsilon) = F$ for any ε, the possibility to cheat is avoided. Were compensation not fixed and dependent on the size of a shock, the shock-affected country would overstate ε if $\frac{\partial F}{\partial \varepsilon} < 0$ and understate ε if $\frac{\partial F}{\partial \varepsilon} > 0$. Thus, asymmetric information requires compensation to be constant to avoid incentive compatibility problems.

A simple way of achieving this is to assume compensation to be monetary and at a pre-determined fixed level. Henceforth, it will thus be assumed that a fixed monetary transfer to the trading partner must be incurred for using the escape clause.

5.2 The First-Best Compensation Cost

The cost F determines the threshold level η_{EC}, below which cooperation is preferred to invoking the escape clause. It is determined as follows

$$\begin{aligned} C \succcurlyeq EC \Leftrightarrow & \; p_C\{W_C^\varepsilon - [W_D^\varepsilon - F]\} + (1-p_C)\{W_S^\varepsilon - [W_N^\varepsilon - F]\} \geq 0 \\ \Leftrightarrow & \; W_C^\varepsilon - W_D^\varepsilon + F \geq 0 \\ \Leftrightarrow & \; F \geq \varepsilon(W_D - W_C) \\ \Leftrightarrow & \; \varepsilon \leq \frac{F}{W_D - W_C} \equiv \eta_{EC}. \end{aligned} \tag{9}$$

Hence, by attaching a fixed compensation cost to the use of the escape clause, the

[28] See also Rosendorff and Milner (2001).

[29] This is reminiscent of what Ethier (2001) describes as the reciprocal-conflict problem, which arises because ex ante, governments know that they might want to deviate, or they might want to retaliate against a deviation by the trading partner.

agreement in fact becomes state-contingent. A country will apply the escape clause if and only if the shock it experiences is sufficiently large. An alternative interpretation is that, by reallocating revenues between the two countries, the compensation cost ensures the truthful revelation of the shock. This interpretation is similar to Feenstra and Lewis (1991), where the transfer of rents from trade restrictions also has an informational role.

It immediately follows from (8) that $\eta_{EC} \leq \overline{\eta}$ holds.[30] Since a higher cost F of exercising the escape clause will constrain its use, the threshold value increases in F. The expected per-period payoff $W_E = (1-\beta)V_C^{EC}$ is given by (for derivation, see the Appendix):

$$W_E = W_N + [W_D - W_N]\Phi(\eta_{EC}) - [W_D - W_C] \int_0^{\eta_{EC}} \varepsilon d\Phi. \qquad (10)$$

The fact that the compensation cost F has no direct effect on the expected per-period outcome is hardly surprising, because ex ante, the probability of using the escape clause and thus incurring the cost for doing so is equal to the probability of being exposed to the use of the escape clause and thus, receiving compensation of the same size. The compensation cost F nevertheless has an indirect effect on the expected per-period outcome via its influence on the threshold value η_{EC}.

Next, the first-best solution for F, i.e. the optimal solution in the absence of the participation constraint given by (8), for any given degree of liberalization is derived.

Proposition 1 *For any strictly positive degree of liberalization, the first-best solution for the threshold value, above which the escape clause is used, is given by*

$$\widehat{\eta}_{EC} = \frac{1}{\tau}, \qquad (11)$$

[30] This might lead to the conclusion that the inclusion of an escape clause leads to a lower probability of cooperation being chosen. Note, however, that $\overline{\eta}$ depends on V_C^{EC} and not on V_C. As will be demonstrated in subsection 5.5, the presence of an escape clause with the maximum possible compensation cost will unambiguously increase the expected per-period payoff and thus, make $\overline{\eta}$ larger than it would be in the absence of an escape clause.

while the first-best solution for the compensation cost is given by

$$\widehat{F} = \widehat{\eta}_{EC}[W_D - W_C] = W_D - W_N. \tag{12}$$

Proof. Given a certain degree of liberalization, the optimality condition for equation (10) is given by

$$\begin{aligned}\frac{\partial W_E}{\partial \eta_{EC}} &= [W_D - W_N]\varphi(\eta_{EC}) - \eta_{EC}\varphi(\eta_{EC})[W_D - W_C] \\ &= \varphi(\eta_{EC})\{[W_D - W_N] - \eta_{EC}[W_D - W_C]\} = 0 \\ &\Leftrightarrow \varphi(\eta_{EC}) = 0 \text{ or } \eta_{EC} = \frac{W_D - W_N}{W_D - W_C} = \frac{1}{\tau}.\end{aligned}$$

Since by assumption $\varphi(\varepsilon) > 0$ for $\varepsilon \in (0, \infty)$, it immediately follows that $\frac{\partial W_E}{\partial \eta_{EC}} \geq 0$ if and only if $\eta_{EC} \leq \frac{1}{\tau}$. Thus, for a certain degree of liberalization, W_E has a unique maximum for $\widehat{\eta}_{EC} = \frac{1}{\tau}$ which, using equation (9), translates into $\widehat{F} = W_D - W_N$ as the first-best value for the compensation cost. ■

The following lemma is straightforward.

Lemma 3 *Increasing the degree of liberalization leads to an increase in the first-best solution for the compensation cost associated with exercising the escape clause and, although using the escape clause becomes costlier, to an increase in its use.*

Proof. Increasing τ implies a decrease in $\widehat{\eta}_{EC} = \frac{1}{\tau}$. Since $W_D - W_N$ unambiguously increases in τ, it follows that $\widehat{F}$ increases in τ. ■

It is the case that $\lim_{\tau \to 0+} \widehat{\eta}_{EC} = \infty$ and that an increase in the degree of liberalization is associated with a rapid fall in the first-best threshold value, above which the escape clause is used.[31] Thus, a higher degree of liberalization will lead to an increase in the use of the escape clause when the first-best compensation cost can

[31] Note that η_{EC}^{opt} is independent of the discount factor.

be applied. Hence, although the incentives to apply the escape clause are inhibited by making its use costlier, the size of the increase in these incentives is sufficiently strong to outweigh the higher cost, thereby making the use of the escape clause more frequent as the degree of liberalization is increased.

The following proposition demonstrates that the first-best compensation cost for using the escape clause is such that it fully incorporates the expected impact of its use on the trading partner.

Proposition 2 *The first-best compensation cost for using the escape clause is such that it equals the expected loss the trading partner incurs by its use, i.e. on average, the trading partner is fully compensated.*

Proof. Consider the difference in expected payoffs between the trading partner applying the escape clause and paying compensation F for doing so and the trading partner making no use of the escape clause. This difference is given by

$$\begin{aligned}
&\int_0^{\eta}[(W_S^{\varepsilon}+F)-W_C^{\varepsilon}]d\Phi+\int_{\eta}^{\infty}[(W_N^{\varepsilon}+F-F)-(W_D^{\varepsilon}-F)]d\Phi\\
&=F-\int_0^{\infty}(W_D^{\varepsilon}-W_N^{\varepsilon})d\Phi\\
&=F-(W_D-W_N).
\end{aligned}$$

Since $\widehat{F}=W_D-W_N$, it immediately follows that the expected impact of the trading partner using the escape clause, while paying compensation, is zero when the first-best compensation cost is applied. ■

Thus, the first-best compensation cost is such that it equals the expected loss from being exposed to the use of the escape clause by the trading partner, i.e., on average, the exposure to the use of the escape clause is fully compensated. Hence, the expected impact of exercising the escape clause is fully incorporated, i.e. it is neutral in the sense of, on average, not adversely affecting the trading partner.

This result highlights the efficiency-enhancing role played by compensation. By incurring a cost for invoking the escape clause, the negative impact of the escape clause is, to some degree, carried by the escape-clause-using country. Thus, the effect of the escape clause is internalized in the decision on whether to apply the escape clause. It immediately follows from the proof of the previous proposition that if the cost for using the escape clause is below the optimal level, the negative impact of the escape clause on the trading partner is only partially internalized. Hence, the escape clause will then too often be used in relation to what is globally optimal.

An important implication of the previous proposition is that it renders support to the principle of reciprocity.[32] If the first-best compensation cost is applied, the escape-clause-exposed country will, on average, be compensated to an extent equivalent to the loss incurred by the escape clause.

It is, however, necessary to emphasize that ex post, it may be the case that a country that has been exposed to the use of the escape clause by its trading partner is adversely affected. If it cooperates, this is the case if $\widehat{F} < W_C^\varepsilon - W_S^\varepsilon$, and if it deviates, this is the case if $\widehat{F} < W_D^\varepsilon - W_N^\varepsilon = W_C^\varepsilon - W_S^\varepsilon$ (by (A2)). Thus, when a country is exposed to the use of the escape clause by its trading partner, it will be undercompensated if $W_D^\varepsilon - W_N^\varepsilon = W_C^\varepsilon - W_S^\varepsilon > \widehat{F}$ and overcompensated if $W_D^\varepsilon - W_N^\varepsilon = W_C^\varepsilon - W_S^\varepsilon < \widehat{F}$; on average, however, the first-best compensation will be such that it exactly offsets the effects of the escape clause. Nevertheless, it is possible that invoking the escape clause turns out not to be socially optimal ex post. Hence, conflicts about the use of the escape clause may arise after its implementation. In the presence of hidden information, that problem is difficult to resolve, because incentives to misrepresent the true loss will arise.[33] Addressing that issue is beyond the scope of this paper, however.

5.3 The Optimal Compensation Cost Scheme

The upper bound for the compensation cost to make the use of the escape clause preferred to deviation serves as a participation constraint, ensuring that deviation is never chosen. The following proposition prescribes the optimal compensation cost scheme, given this participation constraint.

[32] Theoretical support for the principle of reciprocity can be found in e.g. Bagwell and Staiger (2002) and Ethier (2002). Horn and Mavroidis (1999) also point to the fact that reciprocity is a guiding principle for the countermeasures provided in the Dispute Settlement Understanding.

[33] A possibility could be to include a DSP.

Proposition 3 *The optimal cost scheme is such that the escape-clause-exposed country gets compensation equal to the expected loss incurred by the trading partner ($F^{opt} = \widehat{F}$), or, when that is not possible, compensation to the largest possible degree ($F^{opt} = \overline{F}$). Hence, $F^{opt} = \min(\widehat{F}, \overline{F})$.*

Proof. If $\widehat{F} \leq \overline{F}$, then the first-best compensation cost can be implemented and the highest possible per-period outcome achieved. If, however, $\widehat{F} > \overline{F}$, the first-best compensation cost cannot be applied. Since W_E is strictly increasing in η_{EC} for $\eta_{EC} \in (0, \widehat{\eta}_{EC})$ (see proof proposition 1) and hence in F for $F \in (0, \widehat{F})$, it immediately follows that F should be set as close to $\widehat{F}$ as possible, i.e. $F = \overline{F}$. ■

Hence, the compensation cost should be set as high as possible, if the optimal compensation cost cannot be implemented. This result is intuitive, considering the discussion in the previous subsection on the degree to which the compensation cost internalizes the negative expected effect of using the escape clause. A higher compensation cost implies that the overall expected impact of using the escape clause is taken into account to a higher degree, and thus, the escape clause is less extensively used. If the first-best compensation cost can be implemented, the negative expected effect of using the escape clause is fully internalized, and the escape clause will thus be applied only if it is globally optimal. Else, it will be optimal to let the compensation cost be as close as possible to its optimal level, given the participation constraint.

Whether the first-best compensation cost can be implemented depends on the degree of liberalization and the discount factor. The threshold value of the discount factor, above which optimality can be achieved, is determined in the following lemma.

Lemma 4 *The first-best compensation cost can be implemented if and only if*

$$\beta \geq \frac{1}{1 + \Phi(\widehat{\eta}_{EC})[1 - \frac{E(\varepsilon|\varepsilon \leq \widehat{\eta}_{EC})}{\widehat{\eta}_{EC}}]} \equiv \overline{\beta}(\widehat{\eta}_{EC}). \tag{13}$$

Moreover, $\overline{\beta}$ increases strictly in τ.

Proof. See the Appendix. ■

Thus, given a certain degree of liberalization, the first-best compensation cost can be implemented if and only if the discount factor exceeds a threshold value that is a function of the degree of liberalization (since $\widehat{\eta}_{EC} = \frac{1}{\tau}$). The intuition is straightforward. The participation constraint implies that the cost for applying the escape clause cannot be too high in order to avoid deviation being preferred to using the escape clause. The more weight is attributed to future payoffs, reflected in the discount factor, the higher can the cost incurred for using the escape clause be without violating the participation constraint. Hence, for the first-best compensation cost to be implementable, the discount factor must be sufficiently large.

It is easily verified that $\overline{\beta} \in (\frac{1}{2}, 1)$ for $\tau > 0$. Letting τ go to zero, in which case $\widehat{\eta}_{EC}$ approaches infinity, $\overline{\beta}$ converges to $\frac{1}{2}$. Increasing the degree of liberalization, $\overline{\beta}$ increases strictly, converging to unity as τ approaches infinity. The next proposition immediately follows.

Proposition 4 *If $\beta > \frac{1}{2}$, the first-best compensation cost is implementable for degrees of liberalization sufficiently close to zero; for higher degrees of liberalization, this is not the case. If $\beta \leq \frac{1}{2}$, the first-best compensation cost cannot be applied for any strictly positive degree of liberalization.*

Proof. If $\beta \leq \frac{1}{2}$, and hence $\beta < \overline{\beta}$, $\widehat{F}$ cannot be implemented for any $\tau > 0$ and thus, $\overline{F}$ will be implemented. If $\beta > \frac{1}{2}$, it will be the case that $\overline{\beta} \leq \beta$ and thus, $\widehat{F} \leq \overline{F}$ for τ sufficiently close to zero; for higher values of τ, i.e. when $\overline{\beta}(\widehat{\eta}_{EC}) > \beta$, it will, however, be the case that $\widehat{F} > \overline{F}$ and hence, $\overline{F}$ will be implemented. ■

To conclude this subsection, given a certain discount factor and a certain degree of liberalization, the optimal compensation cost and the resulting threshold value, above which the escape clause will be exercised, are given by $F^{opt} = \min(\widehat{F}, \overline{F})$ and $\eta_{EC}^{opt} = \min(\widehat{\eta}_{EC}, \overline{\eta})$.

5.4 The Scope for Liberalization under the Optimal Compensation Cost Scheme

Obviously, finding the degree of liberalization maximizing expected per-period payoffs crucially depends on whether the first-best compensation cost can be implemented. From the previous subsection, we know that the first-best compensation cost cannot be implemented if $\beta \leq \frac{1}{2}$. When $\beta > \frac{1}{2}$, it is implementable for sufficiently low degrees of liberalization, however. First, the case when the maximum compensation cost is applied will be examined and then, the case when the first-best compensation cost can be applied will be assessed.

5.4.1 When the First-Best Solution Cannot Be Implemented ($\widehat{F} > \overline{F}$)

With respect to the threshold value, below which cooperation will be preferred to applying the escape clause, two opposing effects can be identified. On the one hand, by allowing for deviation under the escape clause, the long-term cost of deviation disappears, thus decreasing the likelihood for cooperation being chosen. On the other hand, attaching a cost to invoking the escape clause has a restraining effect on its use, thereby increasing the likelihood for cooperation being preferred. In what follows, the threshold value, below which cooperation will be preferred to applying the escape clause if the first-best cost for using the escape clause cannot be implemented, will be determined.

In this case, the highest possible cost $\overline{F}$ will be implemented. Using equations (9), (8) and (10), an equation similar to that obtained for the case when there is no escape clause (cf. equation (6)) can be obtained for the case where the maximum compensation cost $\overline{F}$ is applied

$$(9): \qquad \eta_{EC} = \frac{\overline{F}}{W_D - W_C} \overset{(8)}{=} \beta \frac{V_C^{EC} - V_D}{W_D - W_C} = \frac{\beta}{1-\beta} \frac{W_E - W_N}{W_D - W_C}$$

$$(10): \qquad W_E - W_N = [W_D - W_N]\Phi(\eta_{EC}) - [W_D - W_C] \int_0^{\eta_{EC}} \varepsilon d\Phi$$

$$\Rightarrow \eta_{EC} = \frac{\beta}{1-\beta} \{ \frac{\Phi(\eta_{EC})}{\tau} - \int_0^{\eta_{EC}} \varepsilon d\Phi \} \equiv f_{EC}(\eta_{EC}) \qquad (14)$$

This equation defines the consistent solutions when the largest possible compensation cost is applied. It turns out that this equation shares the same features as the corresponding equation for the case with no escape clause.

Lemma 5 *There exists at least one stable solution ($\eta_{EC} = 0$) to equation (14).*

Proof. Since $f_{EC}(0) = 0$, it follows that $\eta_{EC} = 0$ is always a solution to (14). ■

Letting η_{EC} go to infinity, f_{EC} converges to $\frac{\beta}{1-\beta}\frac{1-\tau}{\tau}$. There may or may not exist further solutions, depending on the discount factor and the degree of liberalization.

By increasing the discount factor, $f_{EC}(\eta_{EC})$ increases for any given η_{EC}. Since letting β approach unity will make the term $\frac{\beta}{1-\beta}$ very large, it is the case that for a sufficiently high discount factor, there will be values of η_{EC}, for which $f_{EC}(\eta_{EC}) \geq \eta_{EC}$. By decreasing the discount factor, $f_{EC}(\eta_{EC})$ decreases for any given η_{EC} and eventually, $f_{EC}(\eta_{EC}) < \eta_{EC}$ for all $\eta_{EC} > 0$. Thus, there exists a threshold value of β, below which there exist no further solutions for equation (14).

The reasoning is similar for the degree of liberalization. Letting τ approach zero, $f_{EC}(\eta_{EC})$ goes to infinity for any $\eta_{EC} > 0$. As τ increases, $f_{EC}(\eta_{EC})$ decreases for any given η_{EC}. If τ is sufficiently large, $f_{EC}(\eta_{EC}) < \eta_{EC}$ for all $\eta_{EC} > 0$. In this case, there will exist no further solution to equation (14). However, for τ sufficiently close to zero, there will at least be an interval of values for η_{EC}, for which $f_{EC}(\eta_{EC}) \geq \eta_{EC}$.

Define η'_{EC} as the largest solution to (14), i.e. $\eta'_{EC} \equiv \max\{\eta | \eta = f_{EC}(\eta)\}$. The following lemma demonstrates why regarding η'_{EC} as the preferred threshold value is reasonable.

Lemma 6 *η'_{EC} renders the highest continuation value among all solutions to (14).*

Proof. See the Appendix. ■

As above, it will be assumed that both countries share an interest in fostering a belief that η'_{EC}, being the solution that renders the highest expected payoff, is the threshold value applied. Hence, η'_{EC} will be referred to as the relevant solution to (14).

It is easily seen that the expression for f_{EC} is similar to the expression for f, given by (6). An interesting question is how consistent solutions to equations (6) and (14) differ, and what the implications of these differences are. The following proposition addresses this issue.

Proposition 5 *For any strictly positive degree of liberalization, the ex ante likelihood of cooperation being chosen is strictly larger in the presence of an escape clause with the maximum possible compensation cost, than in the absence of an escape clause, as long as there exists a strictly positive solution to (14).*

Proof. It is the case that $f(0) = f_{EC}(0)$. As η goes to infinity, both f and f_{EC} converge to $\frac{\beta}{1-\beta}\frac{1-\tau}{\tau}$ (given that $\beta < 1$ and $\tau > 0$). For any $0 < \eta < \infty$ it is, however, the case that

$$f(\eta) = \frac{\beta\Phi(\eta)}{1-\beta\Phi^2(\eta)}\{\frac{\Phi(\eta)}{\tau} - \int_0^{\eta} \varepsilon d\Phi\}$$

$$< \frac{\beta}{1-\beta}\{\frac{\Phi(\eta)}{\tau} - \int_0^{\eta} \varepsilon d\Phi\} = f_{EC}(\eta).$$

The fact that $f_{EC}(\eta) > f(\eta)$ for any $\eta \in (0, \infty)$ implies that $f_{EC}(\eta') > \eta'$ if $\eta' > 0$. Hence, whenever $\eta'_{EC} > 0$ and $\tau > 0$ (implying $\eta' < \infty$ and $\eta'_{EC} < \infty$), it immediately follows that $\eta'_{EC} > \eta'$ and thus, $\Phi(\eta'_{EC}) > \Phi(\eta')$. ■

This result is somewhat counter-intuitive. The inclusion of an escape clause should make the choice of non-cooperation less costly, since future cooperation is preserved. However, the maximum possible compensation cost equals the expected discounted future gain from preserving the agreement, thus making a country indifferent between applying the escape clause and deviating from the agreement. Thus, the maximum possible compensation cost in fact does not make the choice of non-cooperation less costly, as compared to the case with no escape clause.[34]

It is also important to keep in mind that, while the original purpose of including an escape clause in the agreement was to avoid deviations from the agreement, the attachment of a cost for using it plays the role of taming the frequency of its use which, in turn, has an efficiency-enhancing effect. Making the cost of the escape clause as large as possible, thereby making the degree of internalization of its use as

[34] The fact that the ex ante probability of choosing cooperation is larger in the presence of the escape clause suggests the continuation value to be larger in the presence of an escape clause. In subsection 5.5, it is proven that this is indeed the case.

large as possible, thus inhibits the frequency of its use, and to such an extent that countries actually opt for cooperation to a higher degree than in the absence of the escape clause.

Besides sustaining the agreement and enhancing efficiency, it can be shown that there is a third gain from including an escape clause with an associated compensatory cost at its maximum level. The scope for liberalization increases unambiguously, as demonstrated by the following proposition.

Proposition 6 *For any discount factor $\beta \in (0,1)$ the range of degrees of liberalization, for which the ex ante probability for choosing cooperation is strictly positive (i.e. $\eta'_{EC} > 0$), is strictly larger with an escape clause associated with the maximum possible compensation cost than without any escape clause.*

Proof. Since $f_{EC}(\eta) > f(\eta)$ for any $\eta \in (0, \infty)$, it immediately follows that for any given $\beta \in (0,1)$, the threshold value for τ, above which $\eta'_{EC} = 0$, must be strictly larger than the threshold value for τ, above which $\eta' = 0$. ■

This result resonates well with the often promoted argument that the inclusion of a safeguard facilitates more far-reaching liberalization. An immediate implication of the previous proposition is that there exist degrees of liberalization that are too high to be implementable in the absence of an escape clause, but that can be implemented under an escape clause with the maximum possible compensation cost. Hence, liberalization can be pushed further when an escape clause with the maximum possible compensation cost is included in the agreement. Thus, the inclusion of an escape clause under the optimal compensation cost scheme can be said to lead to freer trade. Whether better outcomes are yielded is addressed in subsection 5.5.

5.4.2 When the First-Best Solution Can Be Implemented ($\widehat{F} \leq \overline{F}$)

From $\widehat{F} \leq \overline{F}$, it immediately follows that $\widehat{\eta}_{EC} < \overline{\eta}_{EC}$. Thus, the implied likelihood of cooperation being chosen will be lower under the first-best compensation cost than under the maximally possible compensation cost. Whether it is lower than in the absence of an escape clause is difficult to determine. The scope for liberalization under the first-best compensation cost is smaller than under the maximum compensation cost, as is demonstrated in the following lemma.

Lemma 7 *The range of degrees of liberalization, in which the first-best compensation cost can be applied, is narrower than the range of degrees of liberalization, in which the maximal compensation cost can be applied.*

Proof. See the Appendix. ■

Thus, it is possible to push liberalization beyond the range where the first-best compensation cost is implementable, as long as the participation constraint is satisfied. Whether it is optimal to push the degree of liberalization beyond the point above which the first-best compensation cost cannot be implemented remains to be assessed.

Analytically, it is very difficult to determine the optimal degree of liberalization without introducing further model specifications. It is, however, possible to derive a general result for the case when the first-best compensation cost can be applied for a wide range of degrees of liberalization (i.e. when the discount factor is high). If the optimal degree of liberalization in the static game without uncertainty lies within this range, the optimal degree of liberalization under the escape clause with the first-best compensation cost will exceed the optimal degree of liberalization in the static game, as demonstrated by the next proposition.

Proposition 7 *If the first-best compensation cost is applicable, the degree of liberalization maximizing the expected per-period payoff is strictly larger than the optimal degree of liberalization in the absence of uncertainty.*

Proof. See the Appendix. ■

The result of proposition 7 is somewhat surprising. How can setting the cooperative tariff level below the static optimal cooperative tariff increase the expected per-period payoff? It must be kept in mind that an increase in τ beyond τ^{opt} leads to a fall in the expected outcome under mutual cooperation and a fall in the expected outcome of being deviated against, but also to an increase in the expected outcome under deviation. Since deviations occur for large realizations of ε, i.e. when the gains from deviating are large, while receiving the sucker's payoff only occurs for low realizations of ε, i.e. when the loss from choosing cooperation is small, it is thus the case that the second effect outweighs the first two effects when the degree of liberalization is pushed above its optimal static value.[35]

5.5 The Gains From Introducing an Escape Clause

By introducing an escape clause, it can be avoided that cooperation breaks down. However, it is not straightforward that expected per-period payoffs will also increase in the presence of an escape clause mechanism. It is important to emphasize that the optimal compensation scheme derived in subsection 5.3 is optimal given the participation constraint (8), i.e. it is optimal among fixed cost schemes, under which deviation will never be chosen. For low compensation costs, it will be the case that including an escape clause yields a lower expected per-period payoff than in the absence of the escape clause. To see why, assume the cost for using the escape clause to be zero. Both countries will then choose to invoke the escape clause in every period, i.e. de facto the situation in the absence of an agreement, with Nash outcomes in every period, is replicated. For sufficiently low degrees of liberalization, having no escape clause will thus be preferable to having one without any cost attached to it. It immediately follows that this must be true at least for sufficiently low compensation cost levels. However, the following proposition states that the inclusion of an escape clause unambiguously increases the expected per-period payoff if the optimal compensation cost scheme is implemented.

Proposition 8 *If the maximum compensation cost is applied, the continuation value of the game exceeds the continuation value in the absence of an escape clause ($V_C^{EC} > V_C$), as long as there exists a strictly positive solution to (14).*

Proof. See the Appendix. ■

[35] If it were indeed the case that liberalization has been pushed beyond its static optimum, because of the presence of an escape clause mechanism with the first-best compensation cost attached to it, it would be understandable that someone with a static perspective would argue that liberalization has gone too far.

Hence, when an escape clause with the maximum compensation cost is applied, the efficiency-enhancing effect is sufficiently large to make expected per-period payoffs larger than in the absence of the escape clause for any strictly positive degree of liberalization, for which the maximum compensation cost is implementable. In the proof of proposition 1, the expected per-period payoff was shown to unambiguously increase in the compensation cost for values below the first-best level and unambiguously decrease for values above the first-best level. The above proposition states that applying the maximum compensation cost level yields an expected per-period payoff in excess of what is obtained in the absence of the escape clause. This is true, notwithstanding if the maximum compensation cost is larger than the first-best compensation cost. The next lemma immediately follows.

Lemma 8 *The discounted expected payoff when there is an escape clause and the first-best compensation cost can be applied is strictly larger than the continuation value in the absence of an escape clause mechanism.*

Proof. Applying the optimal compensation cost yields the highest possible expected per-period payoff. Thus, the result immediately follows from the previous proposition. ■

Since in addition to the benefit from increasing possibilities for liberalization, the efficiency-enhancing effect under the optimal compensation cost scheme is sufficiently strong to increase per-period payoffs in relation to when there is no escape clause, an agreement including an escape clause under the optimal compensation scheme will yield an unambiguously better outcome than when there is no escape clause.

6 Concluding Remarks

The introduction of uncertainty into a repeated Prisoner's Dilemma tariff setting game may cause cooperation to break down in finite time. But by introducing an escape clause mechanism allowing for temporary deviation from cooperation, it is possible to always sustain cooperative behavior, even as shocks increasing the one-period gain from deviating occur. In the presence of hidden information, an escape clause with no cost attached to its use would be invoked all the time. Therefore, a cost must be incurred each time the escape clause is applied.

When the optimal compensation cost scheme is implemented, the expected per-period payoff increases for any given degree of liberalization, as compared to the case when there is no escape clause. Moreover, the scope for liberalization unambiguously increases for any given discount factor. Hence, freer trade can be obtained than in the absence of an escape clause.

Several modifications can be made to the present setting. Instead of assuming complete deviation under the escape clause at a fixed cost, a cost varying in the degree of deviation could be considered. A minor deviation would be less costly than a major one, and a government would then choose the appropriate degree of deviation under such a scheme, which would be implementable despite the presence of hidden information.[36] In fact, it is reasonable to assume that higher expected per-period payoffs are yielded under such a flexible scheme than under the present fixed-cost scheme.

Another interesting extension would be to examine whether the inclusion of an escape clause could be useful even when it is not necessary to sustain agreement-conform behavior, i.e. when cooperation is always preferred to deviation.[37]

It should also be emphasized that it is the expected impact on the trading partner that is internalized by the compensation cost in the present model. However, as pointed out, the exact size of the negative effect on the trading partner depends on the state it is in. Hence, compensation may not be sufficient to offset the negative impact of the escape clause. Allowing for ex post adjustments once an escape clause has been exercised could help alleviate such problems and is an interesting topic for further research.

[36] Such a flexible scheme is more in line with the argument put forward in Bagwell and Staiger (2002), that the enforceable level of cooperation may change with the underlying conditions, suggesting that countries cannot be rigidly held to tariff commitments in a self-enforcing agreement. Thus, flexibility should be provided to adjust tariff levels as underlying circumstances change.

[37] In the present context, this could be the case when there exists an upper bound on the short-term gain from deviating. There is likely to be an efficiency-enhancing effect due to the inclusion of an escape clause.

Most importantly, the basic assumption of instant compensation whenever the escape clause is invoked, necessitates some discussion. A more realistic scenario, in line with the legal framework of the GATT-WTO, described in section 2, would be to have an escape clause associated with no cost during the first period of its use. Thus, a cost would only be incurred if its use were to exceed one period. In the present context of shocks that are independent across time, this modification would yield similar, albeit analytically more intractable, solutions.[38] Depending on whether the escape clause was used in the previous period, two consistent solution equations and hence, two different threshold values would be obtained. The threshold value below which cooperation is preferred to using the escape clause would be lower if the escape clause were not invoked in the previous period than if it were. This is obvious, given that in the former case, no cost is associated with the use of the escape clause. What is crucial is that even if compensation need not be instant, the prospect of having to compensate in the future would have a restrictive effect on the use of the escape clause. A country would have to choose between applying the escape clause at no cost in the current or in some future period. Hence, there would be an incentive to preserve the right to use the escape clause without cost. Applying the escape clause at no cost would therefore be associated with an expected future cost, whereby its use would possibly be restrained.[39] The efficiency-enhancing effect of having a cost attached to the use of the escape clause would thus be present also in this slightly more realistic scenario.

[38] Alternatively, persistence of shocks could be introduced.

[39] Theoretically, it is possible that the escape clause is always invoked if it was not applied in the previous period, while it is more restrictively used if implemented in the previous period.

7 References

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

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

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

Bagwell, Kyle and Staiger, Robert W. (1999), *An Economic Theory of GATT*, American Economic Review, 89(1), 215-248

Bagwell, Kyle and Staiger, Robert W. (2002), *The Economics of the World Trading System*, The MIT Press, Cambridge, Massachusetts

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

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

Ethier, Wilfried J. (2001), *Theoretical Problems in Negotiating Trade Liberalization*, European Journal of Political Economy, 17, 209-232

Ethier, Wilfried J. (2002), *Escape and Entry Mechanisms in the Multilateral Trade System*, PIER Working Paper 02-009, Department of Economics, University of Pennsylvania

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

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

Herzing, Mathias (2005), *Can Self-Destructive Trade Agreements Be Optimal?*, mimeo, Institute for International Economic Studies, Stockholm University, Sweden

Hochman, Gal (2004), *The Interdependence of Safeguard Measures and Tariffs*, mimeo, William Davidson Faculty of Industrial Engineering and Management, Technion, Israel

Hoekman, Bernard M. and Kostecki (1995), Michael M., *The Political Economy of the World Trading System*, Oxford University Press, Oxford, UK

Horn, Henrik and Mavroidis, Petros C. (1999), *Remedies in the WTO Dispute Settlement System and Developing Country Interests*, mimeo, Economic Research and Analysis Division, World Trade Organization, Geneva, Switzerland

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

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

Ludema, Rodney D. (2001), *Optimal International Trade Agreements and Dispute Settlement Procedures*, European Journal of Political Economy, 17(2), 355-376

Maggi,Giovanni (1999), *The Role of Multilateral Institutions in International Trade Cooperation*, American Economic Review, 89(1), 190-214

Martin, Alberto and Vergote, Wouter (2004), *A Case for Strategic Antidumping*, mimeo, Columbia University, New York

Mayer, Wolfgang (1981), *Theoretical Considerations on Negotiated Tariff Adjustments*, Oxford Economic Papers, 33(1), 135-153

Milgrom, Paul R., North, Douglass C., and Weingast, Barry R. (1990), *The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs*, Economics and Politics, 2(1), 1-23

Riezman, Raymond (1982), *Tariff Retaliation from a Strategic Viewpoint*, Southern Economic Journal, 48(3), 583-593

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

Rosendorff, B. Peter and Milner, Helen (2001), *The Optimal Design of International Trade Institutions: Uncertainty and Escape*, International Organization, 55(4), 829-857

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

Schott, Jeffrey J. (1994), *The Uruguay Round: An Assessment*, Institute for International Economics, Washington, DC, USA

World Trade Organization (1994a), *The General Agreement on Tariffs and Trade*, Geneva, Switzerland

World Trade Organization (1994b), *Agreement on Safeguards*, Geneva, Switzerland

8 Appendix

8.1 Derivation of V_C

The continuation value V_C depends on the probability of the trading partner opting for cooperation (p) and all possible realizations of the domestic shock ε.

$$
\begin{aligned}
V_C &= p\{\int_0^\eta [W_C^\varepsilon + \beta V_C]d\Phi + \int_\eta^\infty [W_D^\varepsilon + \beta V_D]d\Phi\} \\
&\quad + (1-p)\{\int_0^\eta [W_S^\varepsilon + \beta V_D]d\Phi + \int_\eta^\infty [W_N^\varepsilon + \beta V_D]d\Phi\} \\
&\overset{(A3)}{=} p\{\int_0^\eta [W_D^\varepsilon - \varepsilon[W_D - W_C] + \beta V_C]d\Phi + \int_\eta^\infty [W_D^\varepsilon + \beta V_D]d\Phi\} \\
&\quad + (1-p)\{\int_0^\eta [W_N^\varepsilon - \varepsilon[W_D - W_C] + \beta V_D]d\Phi + \int_\eta^\infty [W_N^\varepsilon + \beta V_D]d\Phi\} \\
&= p\{\int_0^\infty W_D^\varepsilon d\Phi - \int_0^\eta \varepsilon[W_D - W_C]d\Phi + \int_0^\eta \beta V_C d\Phi + \int_\eta^\infty \beta V_D d\Phi\} \\
&\quad + (1-p)\{\int_0^\infty [W_N^\varepsilon + \beta V_D]d\Phi - \int_0^\eta \varepsilon[W_D - W_C]d\Phi\} \\
&= p\{W_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + \beta p V_C + \beta(1-p)V_D\} \\
&\quad + (1-p)\{W_N + \beta V_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi\} \\
&= pW_D + (1-p)W_N - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + \beta p^2 V_C + \beta(1-p^2)V_D \\
\Leftrightarrow [1-\beta p^2]V_C &= pW_D + (1-p)W_N - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + \beta(1-p^2)V_D \\
&\overset{(3)}{=} pW_D + (1-\beta)(1-p)V_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + \beta(1-p^2)V_D \\
&= pW_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + (1-p)(1+\beta p)V_D \\
\Leftrightarrow V_C &= \frac{1}{1-\beta p^2}\{pW_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi + (1-p)(1+\beta p)V_D\} \\
&= V_D + \frac{1}{1-\beta p^2}\{pW_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi - (1-\beta)pV_D\} \\
&\overset{(3)}{=} V_D + \frac{1}{1-\beta p^2}\{pW_D - [W_D - W_C]\int_0^\eta \varepsilon d\Phi - pW_N\} \\
&= V_D + \frac{1}{1-\beta p^2}\{p[W_D - W_N] - [W_D - W_C]\int_0^\eta \varepsilon d\Phi\} \\
&= V_D + \frac{1}{1-\beta\Phi^2(\eta)}\{[W_D - W_N]\Phi(\eta) - [W_D - W_C]\int_0^\eta \varepsilon d\Phi\} \qquad (4)
\end{aligned}
$$

8.2 Proof of Lemma 2

Let η be a solution to (6), i.e.

$$\eta = f(\eta) = \frac{\beta\Phi^2(\eta)}{1-\beta\Phi^2(\eta)}\{\frac{1}{\tau} - E(\varepsilon|\varepsilon \le \eta)\}$$
$$\Leftrightarrow [W_D - W_N] - [W_D - W_C]E(\varepsilon|\varepsilon \le \eta) = [W_D - W_C]\frac{[1-\beta\Phi^2(\eta)]\eta}{\beta\Phi^2(\eta)}. \qquad (*)$$

The derivative of the continuation value with respect to η is given by

$$\begin{aligned}
\frac{\partial V_C}{\partial \eta} &= \frac{1}{1-\beta\Phi^2(\eta)}\{[W_D - W_N]\varphi(\eta) - [W_D - W_C]\varphi(\eta)\eta\} \\
&\quad + \frac{2\beta\Phi(\eta)\varphi(\eta)}{[1-\beta\Phi^2(\eta)]^2}\{[W_D - W_N]\Phi(\eta) - [W_D - W_C]\int_0^\eta \varepsilon d\Phi\} \\
&= \frac{\varphi(\eta)}{[1-\beta\Phi^2(\eta)]^2}\left\{\begin{array}{c} \{[W_D - W_N] - [W_D - W_C]\eta\}[1-\beta\Phi^2(\eta)] \\ +2\beta\Phi^2(\eta)\{[W_D - W_N] - [W_D - W_C]E(\varepsilon|\varepsilon \le \eta)\} \end{array}\right\} \\
&\overset{(*)}{=} \frac{\varphi(\eta)}{[1-\beta\Phi^2(\eta)]^2}\left\{\begin{array}{c} \{[W_D - W_N] - [W_D - W_C]\eta\}[1-\beta\Phi^2(\eta)] \\ +2\beta\Phi^2(\eta)[W_D - W_C][1-\beta\Phi^2(\eta)]\frac{\eta}{\beta\Phi^2(\eta)} \end{array}\right\} \\
&= \frac{\varphi(\eta)}{1-\beta\Phi^2(\eta)}\{[W_D - W_N] + [W_D - W_C]\eta\} \ge 0.
\end{aligned}$$

Hence, for any solution to (6), it is the case that the continuation value is increasing. It immediately follows that the highest value for η that solves (6) must be that rendering the highest continuation value among the solutions.

8.3 Derivation of W_E

$$\begin{aligned}
W_E &= \Phi(\eta_{EC})\{\int_0^{\eta_{EC}} W_C^\varepsilon d\Phi + \int_{\eta_{EC}}^\infty (W_D^\varepsilon - F)d\Phi\} \\
&\quad + [1-\Phi(\eta_{EC})]\{\int_0^{\eta_{EC}} (W_S^\varepsilon + F)d\Phi + \int_{\eta_{EC}}^\infty W_N^\varepsilon d\Phi\} \\
&= \Phi(\eta_{EC})\{\int_0^\infty W_D^\varepsilon d\Phi - [W_D - W_C]\int_0^{\eta_{EC}} \varepsilon d\Phi - \int_{\eta_{EC}}^\infty F d\Phi\} \\
&\quad + [1-\Phi(\eta_{EC})]\{\int_0^\infty W_N^\varepsilon d\Phi - [W_N - W_S]\int_0^{\eta_{EC}} \varepsilon d\Phi + \int_0^{\eta_{EC}} F d\Phi\}
\end{aligned}$$

$$= \Phi(\eta_{EC})\{W_D - [W_D - W_C]\int_0^{\eta_{EC}} \varepsilon d\Phi - [1-\Phi(\eta_{EC})]F\}$$
$$+ [1-\Phi(\eta_{EC})]\{W_N - [W_D - W_C]\int_0^{\eta_{EC}} \varepsilon d\Phi + \Phi(\eta_{EC})F\}$$
$$= \Phi(\eta_{EC})W_D + [1-\Phi(\eta_{EC})]W_N - [W_D - W_C]\int_0^{\eta_{EC}} \varepsilon d\Phi$$
$$= W_N + \Phi(\eta_{EC})[W_D - W_N] - [W_D - W_C]\int_0^{\eta_{EC}} \varepsilon d\Phi. \qquad (10)$$

8.4 Proof of Lemma 4

(i) The first-best compensation cost is implementable if and only if

$$\widehat{F} \leq \overline{F} \Leftrightarrow W_D - W_N \leq \frac{\beta}{1-\beta}[W_E - W_N]$$
$$\overset{(10)}{=} \frac{\beta}{1-\beta}\{[W_D - W_N]\Phi(\widehat{\eta}_{EC}) - [W_D - W_C]\int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi\}$$
$$\Leftrightarrow W_D - W_N \leq \beta\{[1+\Phi(\widehat{\eta}_{EC})][W_D - W_N] - [W_D - W_C]\int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi\}$$
$$\Leftrightarrow \beta \geq \frac{W_D - W_N}{[1+\Phi(\widehat{\eta}_{EC})][W_D - W_N] - [W_D - W_C]\int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi}$$
$$= \frac{1}{1+\Phi(\widehat{\eta}_{EC}) - \tau\int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi}$$
$$\overset{\tau=\frac{1}{\widehat{\eta}_{EC}}}{=} \frac{1}{1+\Phi(\widehat{\eta}_{EC})[1-\frac{E(\varepsilon|\varepsilon\leq\widehat{\eta}_{EC})}{\widehat{\eta}_{EC}}]} \equiv \overline{\beta} \qquad (13)$$

(ii) It is easily shown that $\overline{\beta}$ decreases unambiguously in $\widehat{\eta}_{EC}$.

$$\frac{d\overline{\beta}}{d\widehat{\eta}_{EC}} = -\frac{\int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi}{\{\widehat{\eta}_{EC}[1+\Phi(\widehat{\eta}_{EC})] - \int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi\}^2} < 0 \qquad \text{(since } \widehat{\eta}_{EC} > 0)$$

Since $\widehat{\eta}_{EC} = \frac{1}{\tau}$, it immediately follows that $\frac{d\overline{\beta}}{d\tau} > 0$.

8.5 Proof of Lemma 6

By showing that for any solution of (14) the expected per-period payoff is increasing in η, it can be concluded that the highest value of η that solves (14) also renders the highest expected per-period payoff. Let η_{EC} be a solution to (14), i.e.

$$\begin{aligned}\eta_{EC} &= \frac{\beta}{1-\beta}\{\frac{\Phi(\eta_{EC})}{\tau} - \int_0^{\eta_{EC}} \varepsilon d\Phi\} \\ &= \frac{\beta}{1-\beta}\Phi(\eta_{EC})\{\frac{1}{\tau} - E(\varepsilon|\varepsilon \leq \eta_{EC})\}\end{aligned}$$

$$\Leftrightarrow [W_D - W_N] - [W_D - W_C]E(\varepsilon|\varepsilon \leq \eta_{EC}) = [W_D - W_C]\frac{(1-\beta)\eta_{EC}}{\beta\Phi(\eta_{EC})}.$$

The derivative of the expected per-period payoff with respect to η_{EC} is given by

$$\begin{aligned}\frac{\partial W_E}{\partial \eta_{EC}} &= [W_D - W_N]\varphi(\eta_{EC}) - [W_D - W_C]\varphi(\eta_{EC})\eta_{EC} \geq 0 \\ &\Leftrightarrow \eta_{EC} \leq \frac{1}{\tau} = \widehat{\eta}_{EC}.\end{aligned}$$

Since it is assumed that the optimal compensation cost cannot be implemented, i.e. $\widehat{F} > \overline{F}$, which implies $\widehat{\eta}_{EC} > \eta_{EC}$, it follows that for any solution to (14), it is the case that the expected per-period payoff is increasing. It immediately follows that the highest value for η_{EC} that solves (14) must be the one rendering the highest expected per-period payoff among the solutions.

8.6 Proof of Lemma 7

The largest possible degree of liberalization $\widehat{\tau}$, at which $\widehat{F}$ can be applied without violating the participation constraint (i.e. when $\widehat{F} = \overline{F}$), is implicitly defined by expression (13)

$$\beta = \overline{\beta}(\frac{1}{\widehat{\tau}}) = \frac{1}{1 + \Phi(\frac{1}{\widehat{\tau}})[1 - \widehat{\tau}E(\varepsilon|\varepsilon \leq \frac{1}{\widehat{\tau}})]}.$$

Since $\frac{1-\beta}{\beta} = \Phi(\frac{1}{\widehat{\tau}})[1 - \widehat{\tau}E(\varepsilon|\varepsilon \leq \frac{1}{\widehat{\tau}})] = \Phi(\frac{1}{\widehat{\tau}}) - \widehat{\tau}\int_0^{\frac{1}{\widehat{\tau}}} \varepsilon d\Phi$, by expression (14)

$f_{EC}(\eta_{EC})$ for $\tau = \widehat{\tau}$ is then given by

$$f_{EC}(\eta_{EC}) = \frac{1}{\widehat{\tau}} \frac{\Phi(\eta_{EC}) - \widehat{\tau} \displaystyle\int_0^{\eta_{EC}} \varepsilon d\Phi}{\Phi(\frac{1}{\widehat{\tau}}) - \widehat{\tau} \displaystyle\int_0^{\frac{1}{\widehat{\tau}}} \varepsilon d\Phi}.$$

It is obviously the case that $f_{EC}(\widehat{\eta}_{EC}) = \frac{1}{\widehat{\tau}} = \widehat{\eta}_{EC}$, i.e. $\widehat{\eta}_{EC}$ solves (14). Taking the first derivative of $f_{EC}(\eta_{EC})$ yields

$$f'_{EC}(\eta_{EC}) = \frac{1}{\widehat{\tau}} \frac{1 - \widehat{\tau}\eta_{EC}}{\Phi(\frac{1}{\widehat{\tau}}) - \widehat{\tau} \displaystyle\int_0^{\frac{1}{\widehat{\tau}}} \varepsilon d\Phi} \varphi(\eta_{EC}).$$

It is easily seen that $f'_{EC}(\eta_{EC}) \geq 0$ if and only if $\eta_{EC} \leq \frac{1}{\widehat{\tau}} = \widehat{\eta}_{EC}$. Hence, $f_{EC}(\eta_{EC})$ has a unique maximum for $\eta_{EC} = \widehat{\eta}_{EC}$. Since $f'_{EC}(\widehat{\eta}_{EC}) = 0$, it immediately follows that $f_{EC}(\eta_{EC}) > \eta_{EC}$ for η_{EC} slightly smaller than $\widehat{\eta}_{EC}$. Hence, τ can be increased beyond $\widehat{\tau}$.

8.7 Proof of Proposition 7

Plugging the optimal compensation cost $\widehat{\eta}_{EC} = \frac{1}{\tau}$ into equation (11) renders the following per-period payoff:

$$W_E = W_N + [W_D - W_N]\Phi(\widehat{\eta}_{EC}) - [W_D - W_C] \int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi.$$

The first derivative of W_E with regard to τ is given by

$$\begin{aligned}
\frac{dW_E}{d\tau} &= \Phi(\widehat{\eta}_{EC})\frac{dW_D}{d\tau} + [W_D - W_N]\varphi(\widehat{\eta}_{EC})\frac{d\widehat{\eta}_{EC}}{d\tau} \\
&\quad - \int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi \frac{d[W_D - W_C]}{d\tau} - [W_D - W_C]\varphi(\widehat{\eta}_{EC})\widehat{\eta}_{EC}\frac{d\widehat{\eta}_{EC}}{d\tau} \\
&= \Phi(\widehat{\eta}_{EC})\frac{dW_D}{d\tau} - \int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi \frac{d[W_D - W_C]}{d\tau} \\
&\quad + [W_D - W_N]\varphi(\widehat{\eta}_{EC})[1 - \tau\widehat{\eta}_{EC}]\frac{d\widehat{\eta}_{EC}}{d\tau}
\end{aligned}$$

$$= \Phi(\widehat{\eta}_{EC})\frac{dW_D}{d\tau} - \int_0^{\widehat{\eta}_{EC}} \varepsilon d\Phi \frac{d[W_D - W_C]}{d\tau}$$
$$= \Phi(\widehat{\eta}_{EC})\{[1 - E(\varepsilon|\varepsilon \leq \widehat{\eta}_{EC})]\frac{dW_D}{d\tau} + E(\varepsilon|\varepsilon \leq \widehat{\eta}_{EC})\frac{dW_C}{d\tau}\}. \quad (15)$$

Since $\frac{dW_D}{d\tau} > 0$ for all τ, it can easily be established that for any $\tau \geq \tau^{opt}$ (i.e. when $\frac{dW_C}{d\tau} \geq 0$) it is the case that $\frac{dW_E}{d\tau} > 0$. In particular, for the optimal static cooperative tariff level τ^{opt} (i.e. when $\frac{dW_C}{d\tau} = 0$), it is the case that $\frac{dW_E}{d\tau} > 0$. Hence, W_E is increasing beyond the static optimal liberalization level.

8.8 Proof of Proposition 8

For very low discount factors, it is the case that $\eta'_{EC} = \eta' = 0$. For a range of higher discount factors, it is the case that $\eta' = 0$ but $\eta'_{EC} > 0$, and the result is trivial. For sufficiently high discount factors, η' is strictly positive, but strictly smaller than η'_{EC}. The gain in expected per-period payoffs under an agreement without an escape clause is in this case given by

$$(1-\beta)V_C - W_N = \frac{1-\beta}{1-\beta\Phi^2(\eta')}\{[W_D - W_N]\Phi(\eta') - [W_D - W_C]\int_0^{\eta'} \varepsilon d\Phi\}$$

$$< [W_D - W_N]\Phi(\eta') - [W_D - W_C]\int_0^{\eta'} \varepsilon d\Phi$$
$$< [W_D - W_N]\Phi(\eta'_{EC}) - [W_D - W_C]\int_0^{\eta'_{EC}} \varepsilon d\Phi \qquad \text{since } \eta' < \eta'_{EC} < \widehat{\eta}_{EC}$$
$$= W_E - W_N$$
$$\Rightarrow V_C < \frac{W_E}{1-\beta} = V_C^{EC}.$$

Thus, the continuation value of the game with an escape clause and the maximum compensation cost for using it is strictly larger than the continuation value of the game without an escape clause.

Chapter 4

Optimal Time Limits on Safeguards in Trade Agreements*

1 Introduction

International agreements in general, and trade agreements in particular, typically include safeguard provisions, which allow a country to (usually temporarily) withdraw a concession made under the agreement in certain contingencies. More specifically, a safeguard makes it possible for a country to introduce protectionist measures, thereby scaling back the agreed-upon liberalization.

This paper addresses the issue of prespecifying the length of the phase, in which a country is permitted to apply a safeguard. The focus is on two issues. First, an analytical framework for how time limits on safeguards are determined is provided and second, the ex post implications of a time limit on the use of safeguards within this framework are examined.

In the present model, there are two symmetric countries with an infinite number of sectors, each of which can be in either of two states. In the good state, there are gains from trade liberalization to be made, while in the bad state, losses will be incurred under trade liberalization. It is in the latter case that invoking a safeguard, allowing for scaling back liberalization, is desirable. Shifts from one state to the

* I am grateful to my advisors Harry Flam and Henrik Horn. Helpful comments from Thomas Eisensee and Per Krusell are appreciated. Thanks also to Christina Lönnblad for editorial assistance. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

other are assumed to be solely stochastically determined by Poisson processes that are identical and independent across countries and between sectors. Hence, there is no upper limit to how long a country can remain in the bad (or the good) state.

Negotiations over a trade agreement between the countries are assumed to cover the provision of a safeguard to be applied during a phase, henceforth referred to as the adjustment phase. The determination of cooperative tariff levels will not be addressed. Hence, the analysis will focus on the optimal choice of length of the adjustment phase for any given cooperative tariffs.

Ex ante, when the agreement is negotiated, a country must weigh two effects of a safeguard against each other. On the one hand, it may wish to implement the safeguard for as long as it deems necessary to stem losses from liberalization. On the other hand, it may find itself exposed to the safeguard measures of another country, in which case it will prefer the safeguard to be applied for as short a period as possible to avoid foregoing the gains from liberalization.

In the negotiations, the two countries are assumed to agree on a rule prespecifying the maximum duration for applying the safeguard. It is shown that what solely determines the optimal length of the adjustment phase is the ratio between the gain from liberalization in the good state and the loss from liberalization in the bad state. If this ratio is smaller than or equal to one, it is optimal to let the adjustment phase be infinite. If it takes on values between one and a threshold value depending on the rate of discounting, letting the length of the adjustment phase be strictly positive, but finite is optimal; the optimal length of the adjustment phase decreases monotonously in this interval of ratios. If the ratio between the gain from liberalization in the good state and the loss from liberalization in the bad state exceeds the threshold value, it is optimal not to allow for an adjustment phase at all, i.e. not to include a safeguard in the agreement. The intuition is straightforward. The larger are the gains from liberalization vis-à-vis the losses, the stronger is the incentive to limit the length of the adjustment phase ex ante.

In the case of a finite adjustment phase, it is shown that the discounted value of future average sector payoffs will eventually start falling. In fact, the time limit on the use of the safeguard will become suboptimal, and its ex post globally optimal value will increase over time. The intuition behind this result is that ex ante, too low a weight is attributed to the situation when a sector has been in the bad state longer than the agreed-upon adjustment phase and hence, would benefit from further protection. The underlying reason is that this situation can only emerge after the agreement has lasted longer than the adjustment phase. In fact, the likelihood of being in that situation increases over time, once the agreement has lasted longer than the adjustment phase. As the likelihood of being in need of extended protection is increasing, the agreed-upon adjustment phase will therefore eventually be too low in relation to what is globally optimal, and the discounted value of future average sector payoffs will be lower than what it would be under the ex post optimal adjustment phase length.

A politically interesting implication of ex post suboptimality is that, in the absence of any readjustment possibilities, the dissatisfaction with the agreement will increase over time, as there is an increase in the share of sectors in the bad state for a period longer than the adjustment phase. What once seemed optimal will, to an increasing degree, be viewed as inappropriate. The pressure to resort to other means of protection might increase, as the share of sectors being in the bad state for a longer period than the adjustment phase increases. Hence, the use of extralegal forms of protection might increase over time. However, while addressing the potential implications of the results derived from the model presented here, resolving them is beyond the scope of the present model.

In a modified version of the model, asymmetry is introduced. For simplicity, asymmetry is assumed to arise because only one sector in one country is exposed to the stochastic changes between states described above. It is demonstrated that if the length of the adjustment phase is negotiated through Nash bargaining, a similar result as previously is obtained. When the losses from liberalization exceed the gains, no upper limit on the safeguard is optimal. If the gains from liberalization exceed the losses by a factor of more than three, not having a safeguard is the optimal solution, while for ratios between gains and losses from liberalization between one and three, it is optimal to have a safeguard with time limits. Hence, including a safeguard with a time limit can be justified also when countries are asymmetric.

Whereas hitherto most contributions to the role of safeguards have assumed strategic interaction to take place in infinitely repeated Prisoner's Dilemma settings[1] , the present paper will apply a different methodological approach. The underlying assumption of infinitely repeated Prisoner's Dilemma games is that retaliation against deviations are, by necessity, always delayed. Such an assumption, implying that it is possible to make short-term gains by deviating against trading partners, is problematic. Criticizing this approach, Ethier (2001) emphasizes that contemporary technology and politics should actually make it possible to instantly punish deviation, thus eliminating the opportunities for short-term gains.[2]

In the present setting, short-term gains are not possible and hence, any deviation can instantly be punished. A government contemplating deviation must thus weigh the immediate response by its trading partner into its decision. If it chooses to deviate, it will do so because it is better off under deviation-cum-retaliation than under mutual cooperation. In the present model, it will be assumed that such situations, under which the strategic interaction is no longer of Prisoner's Dilemma type, may emerge. Hence, while in the aforementioned contributions the role of safeguards is to counter the incentives to deviate in an infinitely repeated Prisoner's Dilemma setting, safeguards are included to alleviate damages from liberalization and time limits on their application serve to strike a balance between winners and losers from liberalization in the present framework.

The next section provides a background to the present paper. Section 3 describes the basic model with two countries and an infinite number of sectors. In section 4, trade liberalization is introduced into this setting, and optimality conditions are derived. In the following section, the ex post implications of a time limit on the use of safeguards are addressed. The asymmetric case is discussed in section 6. Section 7 concludes.

[1] See, for example, Rosendorff and Milner (2001), Herzing (2005b), Hochman (2004) and Martin and Vergote (2004).

[2] While it may be unrealistic to assume that trade agreements serve to solve a Prisoner's Dilemma problem, it may still be the case that a trade agreement introduces Prisoner's Dilemma type of interaction by prescribing reactions to be delayed.

2 Safeguards in Trade Agreements

2.1 Legal Background

A safeguard under a trade agreement is a provision allowing a signatory member to withdraw or cease to apply its normal obligations in order to protect certain overriding interests under specified conditions. Here, the focus will be on the safeguard provisions under the General Agreement on Tariffs and Trade (GATT), with particular emphasis on what is usually referred to as the "escape clause", Article XIX of the GATT. Article XIX of the GATT specifically addresses situations where a country suffers from sudden import surges seriously threatening domestic industries and which may thus be exposed to the temptation to break commitments made under the GATT (see World Trade Organization (1994a)). To avoid deviations from the agreement in such situations, Article XIX §1(a) provides the possibility of temporarily suspending obligations under the agreement to prevent or remedy injury due to liberalization commitments.

Besides the establishment of the World Trade Organization (WTO), the Uruguay Round (1986-1994), among other things, resulted in the Agreement on Safeguards, which contains rules governing the use of safeguard measures, specifically those pursuant to Article XIX (see World Trade Organization (1994b)). The Agreement on Safeguards prescribes safeguard measures to generally be on a Most-Favored-Nation (MFN) basis, although selective applications are permitted (Article 5:2b). Clearly defined time limits on the use of safeguard measures are also specified. Safeguard measures are only permitted for a period not exceeding four years (Art 7:1), except under special circumstances (Art 7:2); the total period of application shall not exceed eight years, however (Art 7:3).[3] Having applied a safeguard measure, it cannot be reinvoked "for a period of time equal to that during which such measure had been previously applied, provided that the period of non-application is at least two years" (Art 7:5). However, an exception is made if the safeguard measure has a duration of 180 days or less, at least one year has elapsed since the introduction of the initial measure and such a measure has not been applied more than twice during five years

[3] Developing countries are granted the possibility of extending a safeguard measure up to ten years, inclusive of extensions (Art 9:2).

preceding its date of introduction (Art 7:6).

2.2 The Economic Rationale for Safeguards

The starting point for the analysis of the inclusion of safeguards in international agreements on cooperation between countries is the observation that there may be ex post incentives to make adjustments to the commitments made under such an agreement. Such incentives may arise due to unforeseen events making the outcome under the ex ante agreed-upon cooperative regime suboptimal from a single country's point-of view. In the context of a trade agreement, the ex ante negotiated degree of liberalization may turn out to be ex post suboptimal for an individual country. In the absence of any adjustment instrument, the incentives to breach the agreement may become sufficiently large to make the cooperative regime unsustainable.

The inclusion of a safeguard (or any other flexibility-enhancing instrument) can be justified by ex ante uncertainty about what contingencies may arise ex post.[4] Ethier (2002) argues that by necessity, trade agreements are incomplete contracts negotiated under uncertainty about the future.[5] Since ex post commitments under an agreement may become politically untenable, the inclusion of a possibility to escape from these commitments to sustain the agreement may be essential. However, a country incurring a loss due to the withdrawal of concessions under the escape clause by its trading partner will also wish to withdraw concessions to compensate the loss. It is shown that it is ex ante optimal to allow a country exposed to the escape clause by its trading partner to withdraw equivalent concessions, thus rendering support to the principle of reciprocity to be applied also with respect to the countermeasures permitted in response to the use of the escape clause.

Bagwell and Staiger (2002) provide justification for reciprocity, in combination with the MFN clause, as an optimal governing principle, not only in negotiations about liberalization, but also with respect to how ex post adjustments to the commitments should be conducted.[6] While the principle of reciprocity serves to neu-

[4] Several contributions have introduced various types of uncertainty into models of strategic interaction between trading countries, e.g. Feenstra (1987), Bagwell and Staiger (1990), Jensen and Thursby (1990), Feenstra and Lewis (1991), Riezman (1991) and Herzing (2005a).

[5] See also Ethier (2001).

[6] See also Bagwell and Staiger (1999).

tralize the impact of tariff setting on world prices and thus, the terms of trade, the MFN principle ensures that the externality associated with tariff setting only travels through world prices and not through local prices. Together, these two principles can thus guide governments towards efficiency in multilateral trade agreements.

Rosendorff and Milner (2001) specifically address the role of an escape clause in a two-country setting where the political pressure for protection at home (and/or more open markets abroad) is subject to a shock unobservable to the trading partner. Since in an infinitely repeated Prisoner's Dilemma cooperation may break down when this pressure is sufficiently high to outweigh the loss incurred through punishment, the possibility of exercising an escape clause, allowing deviation for one period at a fixed cost, is introduced. The cost incurred through the use of the escape clause serves as a signal for the willingness to maintain cooperation in the future. Thus, when political leaders cannot foresee the extent of future domestic demands for more protection at home (and/or more open markets abroad), escape clauses provide the flexibility that allows them to accept a trade agreement.

In a similar approach, Herzing (2005b) also lets countries strategically interact in an infinitely repeated Prisoner's Dilemma tariff setting game under hidden information. Unanticipated shocks that are non-observable to the trading partner influence the incentive to deviate from cooperation. To counter these incentives, an escape clause is included in the agreement, allowing a shock-affected country to temporarily deviate from cooperation without causing infinite reversion to the suboptimal Nash outcome. However, a cost for using the escape clause must be incurred to avoid its permanent use. Differing from Rosendorff and Milner (2001), this cost is not sunk but translates into a benefit for the trading partner, thus alleviating its damage from being exposed to the escape clause. It is shown that under the optimal fixed compensation cost scheme, the expected per-period payoffs increase relative to the case when there is no escape clause. Furthermore, the scope for liberalization increases unambiguously in the presence of an escape clause with the optimal fixed compensation cost.

The need for the possibility of ex post adjustments in the presence of uncertainty is also emphasized in Martin and Vergote (2004), who also let strategic interaction between two countries take place in the presence of hidden information. The antidumping instrument is incorporated to allow for adjustment to changes in political preferences that cannot be observed by the trading partner. Hochman (2004) addresses the issue of whether safeguards should be selectively applied under an agreement between three countries in the presence of commonly observable shocks.

3 The Symmetric Case

3.1 The Setup

There are two countries with an infinite number of symmetric sectors, where each sector is engaged in trade with the corresponding sector in the other country. The assumption of an infinite number of sectors is made for analytical convenience. Due to symmetry, it suffices to focus on one country only.

Time is taken to be continuous. Let the instant payoff of the home country government be the sum of payoffs w^i generated in each sector i: $\Sigma_i w^i$. Each sector i is assumed to be in either of two states, i.e. $\varepsilon^i \in \{\underline{\varepsilon}, \overline{\varepsilon}\}$, where ε^i is stochastically determined and perfectly observable. Let t^i and t^{i*} be the home and the foreign country's tariffs in sector i, respectively. For tractability, the following assumption will be made.

Assumption 1 w^i is independent of all stochastic variables, except ε^i, for any given pair of sector tariffs (t^i, t^{i*}).

This simplifying assumption has several strong implications. First, it implies that there are no externalities across sectors. With an infinite number of sectors, the overall impact of other sectors can reasonably be assumed to be neutral, however. Second, it implies that the state of the corresponding sector in the other country has no impact on domestic sector payoffs. Allowing for positive (or negative) externalities will not qualitatively alter the results obtained.[7]

Let the payoff generated in sector i be defined as $w^i \equiv w^i(\varepsilon^i, t^i, t^{i*})$. For any pair of t^i and t^{i*}, it is the case that $w^i(\overline{\varepsilon}, t^i, t^{i*}) > w^i(\underline{\varepsilon}, t^i, t^{i*})$. Thus, the realization of ε^i

[7] Assuming the number of countries to be very large, in which case the aggregate impact of foreign realizations of the stochastic variable would be constant could, for example, justify assumption 1. A large number of countries would, however, complicate the analysis in other respects.

can, for example, be regarded as reflecting high/low domestic demand or high/low productivity. Henceforth, $\overline{\varepsilon}$ ($\underline{\varepsilon}$) will be referred to as the good (bad) state.

Switches between states (in both directions) are assumed to be governed by Poisson processes that are identical and independent between sectors and across countries. Hence, whenever a sector is in one state, the likelihood p of a switch of states at time T in the future is given by $\rho(T) = 1 - e^{-T}$.[8] The likelihood of a switch is thus independent of how long the country has already been in a state. Theoretically, it is possible that a sector will remain in a state for any finite length of time.

When the game starts, half of the sectors are assumed to be in the bad state[9] , while it is random which sectors are actually in that state. Then, each sector will find itself in either state for varying lengths of time, depending on the stochastic process. An infinite number of sectors implies that half the sectors will be in the good (bad) state at any point in time. Thus, both countries will always be equally well off, and will never have diverging interests.

In the absence of a trade agreement, both countries will in each sector apply their optimal tariff t_N vis-à-vis each other which, for simplicity, is assumed to be trade-inhibiting. Due to symmetry, it is straightforward that the optimal tariff is the same in all sectors in both countries. The home country's current payoff from sector i in the absence of any trade cooperation is given by $w_N(\underline{\varepsilon})$ or $w_N(\overline{\varepsilon})$, depending on in which state it is. The assumption of t_N being trade-inhibiting is made to avoid that t_N is state-dependent and hence, that externalities across countries in corresponding sectors arise.[10]

Let $\overline{w}(\tau)$ be the average sector payoff at time τ.[11] In the absence of trade

[8] The use of continuous rather than discrete time is convenient when calculating the length of the optimal adjustment phase, although there is no qualitative difference between the results obtained under these two different approaches.

[9] Since the likelihood of switching states is the same in either state, this assumption is necessary to keep the shares of sectors in the good (bad) state constant over time.

[10] If t_N and t_N^* were instead state-dependent, i.e. $t_N = t_N(\varepsilon)$ and $t_N^* = t_N^*(\varepsilon^*)$, the Nash payoffs would not only depend on the domestic state, but also indirectly, through the foreign tariff, on the state of the trading partner. Having an infinite number of sectors, however, $w_N(\underline{\varepsilon})$ and $w_N(\overline{\varepsilon})$ could in this case be interpreted as the average payoffs in the respective state. Or, alternatively, $w_N(\underline{\varepsilon})$ and $w_N(\overline{\varepsilon})$ could be seen as expected values.

[11] With a finite number of sectors, this would be the expected payoff at time τ.

cooperation, it is thus the case that

$$\overline{w}_N(\tau) = \frac{1}{2}[w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})].$$

The common rate of discounting future payoffs is given by $\delta \in (0, 1)$, thereby implying a discount factor of $1-\delta$. Let $\overline{v}(\tau)$ be the average discounted flow of sector payoffs at time τ, which in the absence of any trade agreement is given by

$$\overline{v}_N(\tau) = \frac{1}{\delta}\overline{w}_N(\tau) = \frac{1}{2\delta}[w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})].$$

Next, trade liberalization will be introduced in this setting.

3.2 Introducing Liberalization and Escape

The two countries may agree to lower trade barriers, i.e. to agree upon a tariff $t_C < t_N$.[12] Let $w_C(\overline{\varepsilon})$ and $w_C(\underline{\varepsilon})$ be the payoffs generated by a sector in states $\overline{\varepsilon}$ and $\underline{\varepsilon}$, respectively, under a commonly agreed-upon cooperative tariff. The following assumption is crucial.

Assumption 2

$$w_C(\overline{\varepsilon}) > w_N(\overline{\varepsilon})$$

$$w_C(\underline{\varepsilon}) < w_N(\underline{\varepsilon}).$$

This assumption implies that a sector will be better off under liberalization only if it finds itself in state $\overline{\varepsilon}$; being in state $\underline{\varepsilon}$, it will actually be worse off than in the absence of liberalization. This assumption can, for example, be justified if the realization of state $\underline{\varepsilon}$ corresponds to the firms in the sector being uncompetitive, while state $\overline{\varepsilon}$ corresponds to their being competitive. In the former case, a country's sector will lose from liberalization while, in the latter case, it will benefit from lower trade barriers.

It will be assumed that there are no short-term gains to be made, in accordance with, for example, Ethier (2002). In other words, any deviation will be followed by instant retaliation. Hence, interaction is not of Prisoner's Dilemma type.

[12] Due to symmetry, trade liberalization will also be symmetric and hence, the cooperative tariff applies to all sectors in both countries.

Since a sector will be worse off under liberalization, it may, ex ante, when an agreement on liberalization is negotiated, be the case that both countries wish to include a safeguard, allowing for scaling back liberalization in a bad-state sector for a period of time of length λ. On the one hand, a government will want the safeguard to be applied as long as necessary, i.e. until a bad-state sector switches to state $\overline{\varepsilon}$, in which case it will prefer the agreed-upon degree of liberalization for that sector. On the other hand, a government will wish λ to be low, in case a sector is in state $\overline{\varepsilon}$, because good-state sectors unambiguously gain from liberalization.[13]

It must be emphasized that agreements with a predefined length of the adjustment phase constitute a subclass of all possible agreements. Within this subclass, there is a clear-cut rule prescribing exactly for how long a safeguard can be applied, whenever a switch to the bad state has occurred. Applying the safeguard under other conditions, i.e. after the adjustment phase has elapsed, is thus explicitly forbidden and regarded as a breach of the agreement. Alternatively, agreements with more flexible rules governing the use of safeguards could be considered.[14] However, there are strong reasons to assume that the approach taken here, with strictly defined time limits for the use of safeguards under any circumstances, is relevant. After all, the GATT-WTO is a rules-based concept. In fact, as pointed out in the previous section, there are explicit time limits for the use of safeguards under the GATT-WTO. But also from a more practical-realistic point of view, a clear-cut rule, to which participants are required to adhere, may be preferable to a system allowing for more flexibility. In the latter case, there is a danger either that the length of application of the safeguard would be left to countries' discretion, or that coordination across countries would be required (something that might be associated with some cost or be politically unfeasible). While addressing the potential suboptimality of the derived optimal solution under a clear-cut rule, the present paper abstracts away from other subclasses of safeguard agreements.

For tractability, the following assumption will be made.

---

[13] This is reminiscent of the reciprocal-conflict problem, addressed by Ethier (2001), which arises because of conflicting interests with regard to the degrees of punishment to be allowed under a trade agreement.

[14] The optimal solution in the subclass under consideration may actually not be first-best. This will be more thoroughly discussed in the next section.

Assumption 3 Whenever the safeguard is invoked by one country in one sector, both countries revert to the Nash equilibrium in that sector.

There are two important aspects of this assumption. First, it implies that applying the safeguard is equivalent to entirely scaling back liberalization. Naturally, this is a strong assumption. In reality, by imposing protection under a safeguard, countries only revert liberalization to some extent, depending on how large an injury an industry is perceived to suffer.[15] However, it is beyond the scope of this paper to take such considerations into account, because it would require a more specific analytical model.

Furthermore, assumption 3 implies that scaling back liberalization will be mutual rather than unilateral. Hence, a country scaling back liberalization in one sector will also face such a scale-back in the corresponding sector of the other country. Alternatively, and possibly more realistically with regard to the legal provisions governing the escape from commitments under the GATT (see World Trade Organization (1994a)), it could be assumed that the realization of $\underline{\varepsilon}$ leads to unilateral deviation for a maximum duration of λ. Making the assumption that both countries will actually deviate does not qualitatively alter the analysis, however.[16,17]

Thus, it will be the case that, once one sector finds itself in state $\underline{\varepsilon}$, both countries will entirely scale back liberalization for a maximum duration of λ. After this time interval has elapsed, or if state $\overline{\varepsilon}$ is reached before that, both countries will revert to the agreed-upon degree of liberalization in that sector.

It is implicitly assumed that a safeguard can be applied as soon as a sector switches to state $\underline{\varepsilon}$, irrespective of how much time has elapsed since it was previously

[15] Since it is assumed in the present model that the injury in the bad state exactly equals the concession made under liberalization, assuming a complete withdrawal of liberalization is in accordance with Article XIX §1(a) of the GATT (see subsection 2.1).

[16] Under the present assumption of perfect observability, the legitimate use of the safeguard is guaranteed, because any attempt to make a gain by applying the safeguard in the good state can be punished. In the presence of hidden information, however, assumption 3 could be justified as a means of providing the correct incentives for using the escape clause.

[17] What is of importance is that the safeguard-applying country benefits, while its trading partner loses as long as the safeguard is implemented. Assuming that both countries in fact deviate when a safeguard is applied has the same effect on payoffs, although not on the size of these effects. The safeguard-applying country gains, but to a lower degree than under unilateral deviation, while the trading partner loses, albeit less so than under unilateral deviation.

used. As pointed out in subsection 2.1, there are clearly defined limits for when a safeguard measure can be reinvoked. For tractability, this legal restriction on the possibility to apply a safeguard will not be considered here.

4 Optimal Time Limits on Protection under Trade Liberalization

4.1 Implementing a Safeguard

As emphasized in the previous section, an infinite number of sectors implies that both countries will always have identical aggregate payoffs, although sector payoffs may differ within and between the two countries. Hence, the winners and losers from liberalization will be sectors within a country rather than entire countries, and conflicting interests may therefore arise within rather than between countries.

Once the agreement is in place, the average sector payoff will depend on how large a share of the sectors actually apply the safeguard. Since there is an infinite number of sectors, half of all sectors will be in the bad state at any point in time, but not all the bad-state sectors will eventually be allowed to apply the safeguard if $\lambda < \infty$. Those sectors that have been in state $\underline{\varepsilon}$ for a period of time exceeding λ will no longer be protected and hence, be worse off than in the absence of trade liberalization.

Let τ be the time since the agreement was implemented. Let $\mu(\underline{\varepsilon}^{\lambda-}, \tau)$ be the share of sectors having been in state $\underline{\varepsilon}$ for a period shorter than λ, $\mu(\underline{\varepsilon}^{\lambda+}, \tau)$ the share of sectors having been in state $\underline{\varepsilon}$ for a period longer than λ, and $\mu(\overline{\varepsilon}, \tau)$ the share of sectors in state $\overline{\varepsilon}$. These shares are then given by

$$\begin{aligned}
\mu(\underline{\varepsilon}^{\lambda-}, \tau) &= \begin{cases} \frac{1}{2} & \text{if } \tau \leq \lambda \\ \frac{1}{2}\frac{1-e^{-\lambda}}{1-e^{-\tau}} & \text{if } \tau > \lambda \end{cases} \\
\mu(\underline{\varepsilon}^{\lambda+}, \tau) &= \begin{cases} 0 & \text{if } \tau \leq \lambda \\ \frac{1}{2}\frac{e^{-\lambda}-e^{-\tau}}{1-e^{-\tau}} & \text{if } \tau > \lambda \end{cases} \\
\mu(\overline{\varepsilon}, \tau) &= \frac{1}{2}.
\end{aligned}$$

It is easily seen that if $\lambda = 0$, then $\mu(\underline{\varepsilon}^{\lambda-}, \tau) = 0$ and $\mu(\underline{\varepsilon}^{\lambda+}, \tau) = \frac{1}{2}$, and if $\lambda \to \infty$, then $\mu(\underline{\varepsilon}^{\lambda-}, \tau) = \frac{1}{2}$ and $\mu(\underline{\varepsilon}^{\lambda+}, \tau) = 0$. For any $\lambda \in (0, \infty)$, it is, however, the case that the longer an agreement has existed, the lower will the share of sectors having been in state $\underline{\varepsilon}$ for a period shorter than λ be, and the larger will the share of sectors having been in state $\underline{\varepsilon}$ longer than λ be. Letting $\tau \to \infty$, $\mu(\underline{\varepsilon}^{\lambda-}, \tau)$ and $\mu(\underline{\varepsilon}^{\lambda+}, \tau)$ will converge to $\frac{1-e^{-\lambda}}{2}$ and $\frac{e^{-\lambda}}{2}$, respectively. These changes in $\mu(\underline{\varepsilon}^{\lambda-}, \tau)$ and $\mu(\underline{\varepsilon}^{\lambda+}, \tau)$ will have two effects. On the one hand, the risk of being exposed to the situation where further protection would be desirable but is no longer possible (i.e. being in state $\underline{\varepsilon}^{\lambda+}$), increases. On the other hand, the likelihood of being in the position of being subject to a trading partner's protection although this is undesirable (i.e. being in state $\overline{\varepsilon}$ and exposed to the trading partner's corresponding sector being in state $\underline{\varepsilon}^{\lambda-}$) decreases.

Define

$$\Pi \equiv \frac{w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})}{w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})}.$$

Being the ratio between gains and losses from liberalization in the two states, Π can be regarded as a measure of the expected relative benefit from liberalizing trade. In fact, Π can be seen as an indicator of whether liberalization is worthwhile in the absence of any safeguard provisions. If $\Pi < 1$, the losses from liberalization in the bad state exceed the gains in the good state and hence, no liberalization is preferable to liberalization. And if $\Pi > 1$, liberalization will be beneficial overall, even if no safeguards are provided, because the gains from liberalization in the good state are larger than the losses from liberalization in the bad state.

The average sector payoff $\overline{w}$ at time τ for a given λ is given by

$$\begin{aligned}\overline{w}(\tau, \lambda) = {} & \mu(\overline{\varepsilon}, \tau)\{\mu(\underline{\varepsilon}^{\lambda-}, \tau) w_N(\overline{\varepsilon}) + [1 - \mu(\underline{\varepsilon}^{\lambda-}, \tau)] w_C(\overline{\varepsilon})\} \\ & + \mu(\underline{\varepsilon}^{\lambda-}, \tau) w_N(\underline{\varepsilon}) \\ & + \mu(\underline{\varepsilon}^{\lambda+}, \tau)\{\mu(\underline{\varepsilon}^{\lambda-}, \tau) w_N(\underline{\varepsilon}) + [1 - \mu(\underline{\varepsilon}^{\lambda-}, \tau)] w_C(\underline{\varepsilon})\}\end{aligned}$$

Plugging in the value for μ derived above, $\overline{w}(\tau,\lambda)$ can be expressed as follows

$$\overline{w}(\tau,\lambda)=\begin{cases}\frac{1}{2}w_N(\underline{\varepsilon})+\frac{1}{4}w_N(\overline{\varepsilon})+\frac{1}{4}w_C(\overline{\varepsilon}) & \text{if } \tau<\lambda \\ \\ \frac{1}{2}w_C(\underline{\varepsilon})+\frac{1}{2}w_C(\overline{\varepsilon}) & \\ +\frac{1-e^{-\lambda}}{4(1-e^{-\tau})}(3-\frac{1-e^{-\lambda}}{1-e^{-\tau}}-\Pi)[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] & \text{if } \tau\geq\lambda\end{cases}. \quad (1)$$

It immediately follows that

$$\overline{w}(\tau,0)=\frac{1}{2}w_C(\underline{\varepsilon})+\frac{1}{2}w_C(\overline{\varepsilon})$$
$$\lim_{\lambda\to\infty}\overline{w}(\tau,\lambda)=\frac{1}{2}w_N(\underline{\varepsilon})+\frac{1}{4}w_N(\overline{\varepsilon})+\frac{1}{4}w_C(\overline{\varepsilon}).$$

Hence, average sector payoffs will be constant over time in the absence of a safeguard or when the safeguard can be indefinitely applied.

For a given λ, the average discounted flow of future payoffs $\overline{v}$ at time τ is obtained as follows

$$\overline{v}(\tau,\lambda)=\int_0^\infty\overline{w}(\tau+s,\lambda)e^{-\delta s}ds.$$

Since $\overline{w}$ is constant over time when $\lambda=0$ or $\lambda=\infty$, the same is true for the average discounted flow of future payoffs, which in these cases becomes

$$\overline{v}(\tau,0)=\frac{1}{2\delta}w_C(\underline{\varepsilon})+\frac{1}{2\delta}w_C(\overline{\varepsilon})$$
$$\lim_{\lambda\to\infty}\overline{v}(\tau,\lambda)=\frac{1}{2\delta}w_N(\underline{\varepsilon})+\frac{1}{4\delta}w_N(\overline{\varepsilon})+\frac{1}{4\delta}w_C(\overline{\varepsilon}).$$

For $\lambda\in(0,\infty)$ it is, however, the case that $\overline{w}$ changes for $\tau>\lambda$. The following expression for $\overline{v}(\tau,\lambda)$ can be derived (see the Appendix).

$$\overline{v}(\tau,\lambda)=\begin{cases}\frac{1}{4\delta}[2w_N(\underline{\varepsilon})+w_N(\overline{\varepsilon})+w_C(\overline{\varepsilon})] & \\ +\frac{(\Pi-1-e^{-\lambda})e^{-(1+\delta)\lambda}e^{\delta\tau}}{4\delta[1+\delta(1-e^{-\lambda})]}[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] & \text{if } \tau<\lambda \\ \\ \frac{1}{2\delta}[w_C(\underline{\varepsilon})+w_C(\overline{\varepsilon})] & \\ +\frac{1-e^{-\lambda}}{4\delta[1+\delta(1-e^{-\tau})]}[2+3\delta+e^{-\lambda}-\delta\frac{1-e^{-\lambda}}{1-e^{-\tau}} & \\ \quad-(1+\delta)\Pi][w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] & \text{if } \tau\geq\lambda\end{cases}. \quad (2)$$

Next, the optimal choice of adjustment phase length will be determined.

4.2 Optimization

Ex ante, given a degree of trade liberalization, a government will choose the adjustment phase length maximizing the average discounted flow of future sector payoffs at $\tau = 0$, which by (2) is given by

$$\overline{v}(0,\lambda) = \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{(\Pi - 1 - e^{-\lambda})e^{-(1+\delta)\lambda}}{4\delta[1+\delta(1-e^{-\lambda})]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})].$$

Let $\widehat{\lambda} \equiv \arg\max_{\lambda\in[0,\infty]} \overline{v}(0,\lambda)$ be the ex ante optimal adjustment phase. The following proposition relates the optimal choice of adjustment phase length to the relative benefit from introducing trade liberalization.

Proposition 1 *The optimal adjustment phase length unambiguously decreases in the ratio between the gains and losses from liberalization. More specifically*

$$\widehat{\lambda} = \begin{cases} \infty & \text{if } \Pi \leq 1 \\ \ln[\frac{2\delta(1+\delta)}{2+3\delta+\Pi\delta^2 - \sqrt{(2+3\delta+\Pi\delta^2)^2 - 4(\Pi-1)\delta(1+\delta)^3}}] & \text{if } \Pi \in (1, 2+\frac{1}{1+2\delta}) \\ 0 & \text{if } \Pi \geq 2 + \frac{1}{1+2\delta} \end{cases} . \quad (3)$$

Proof. See the Appendix. ■

The intuition behind this result is straightforward. The larger is the ratio of gains to losses from liberalization, the less should liberalization be inhibited. Hence, while not constraining the use of the safeguard is optimal if the gains are outweighed by the losses from liberalization ($\Pi \leq 1$)[18] , not having any safeguard provision is

[18] As previously noted, no liberalization would actually be preferred to liberalization in the absence of the safeguard if $\Pi \leq 1$.

optimal if the gains are sufficiently larger than the losses from liberalization ($\Pi \geq 2 + \frac{1}{1+2\delta}$). For intermediate values of the ratio of gains to losses from liberalization ($\Pi \in (1, 2 + \frac{1}{1+2\delta})$), a safeguard with a time limit is optimal.[19]

It is also important to stress that this result does not hinge on political economic or moral hazard considerations, which might potentially add further incentives for limiting the length of the adjustment phase. An industry in decay (in the bad state) enjoying the protection of a safeguard should have low incentives to take measures to become more competitive (reach the good state). In fact, it might rather lobby for continued protection than take any such steps. Hence, a government might prefer to have an outside commitment by means of a trade agreement, which limits its possibilities of delivering protection. Letting switches between states be exogenously determined, the present model does not take such considerations into account, however.

It is important to emphasize that the derived results hinge on the assumption that ex ante, in negotiations about a trade agreement, the two countries decide to adopt a clear-cut rule governing for how long a safeguard can be applied, and then adhere to it. Thus, it is implicitly assumed that mutually agreed-upon breaches against this rule will not take place, once the agreement has been implemented. Hence, the derived solutions are optimal, subject to the constraint that such a rule is applied. In fact, the derived solutions may not be first-best due to this constraint (see next subsection).

The following lemma demonstrates that the optimal length of the adjustment phase decreases in the discounting rate.

Lemma 1 *If $\widehat{\lambda} \in (0, \infty)$, then $\widehat{\lambda}$ decreases in δ. In particular, $\widehat{\lambda}$ may become zero as δ is increased.*

Proof. See the Appendix. ■

Thus, the more future payoffs are discounted, the shorter the optimal length of the adjustment phase will become, and the lower the threshold value for the ratio between gains and losses from liberalization, above which no safeguard is the optimal

[19] A more thorough interpretation of this result can be found in the next subsection.

solution, will be. Intuitively, this result is not straightforward. Actually, it is not clear why the discount rate should have any impact at all. What drives this result, however, is the fact that a higher discount rate implies that the prospect of having a sector in state $\underline{\varepsilon}^{\lambda+}$ sometime in the future carries less weight. Hence, the negative impact of a shorter adjustment phase on this category of sectors is given less weight, thereby moving the optimal adjustment phase length solution in favor of what is optimal for sectors in state $\overline{\varepsilon}$.

From an analytical point of view, the case when future payoffs carry the same weight as current payoffs is of interest.[20] Letting δ approach zero, $\widehat{\lambda} = \ln(\frac{2}{\Pi-1})$ for $\Pi \in (1, 3)$. Hence, the qualitative result in proposition 1, with finite solutions for $\widehat{\lambda}$ for intermediate values of Π, is valid even as the weight given to future payoffs approximates the weight given to current payoffs.

4.3 Ex ante Suboptimality

As previously emphasized, having a rule that is supposed to apply in all contingencies may act as a constraint. Optimality under such a rule may not yield the first-best solution, which is confirmed by the following proposition.

Proposition 2 *The optimal solution given by (3) is first-best if and only if* $\Pi \leq 1$.

Proof. Consider any point in time after the agreement has been implemented. There will exist four categories of sectors, depending on the own state and the state of the corresponding sector in the other country. The first category comprises corresponding sectors in state $\overline{\varepsilon}$ in both countries, in which case no safeguard is applied. Thus, this category need not be considered. If the corresponding sectors in the two countries are both in state $\underline{\varepsilon}$, both are in need of the safeguard, irrespective of how long they have been in that state. Hence, no time limit ($\lambda = \infty$) on the use of the safeguard is optimal for this category. The remaining two categories comprise corresponding sectors in different states in the two countries. It is straightforward

[20] Since the value of discounted future flows of payoffs would be infinitely large for any λ in this case, it is not possible to analytically solve the optimization problem. Letting the discount rate go to zero nevertheless renders insights into how the optimal adjustment phase length is affected, if future payoffs are given approximately the same weight as present payoffs.

that applying the safeguard is globally optimal if $\Pi < 1$, while not applying it is globally optimal if $\Pi > 1$ (and aggregate sector payoffs are the same with or without a safeguard, if $\Pi = 1$). Thus, for these two categories, no time limit ($\lambda = \infty$) on the use of the safeguard is optimal if $\Pi < 1$, while not allowing the safeguard to be applied at all ($\lambda = 0$) is optimal if $\Pi > 1$ (and any $\lambda \in [0, \infty]$ is optimal if $\Pi = 1$). It immediately follows that under a rule applied under all contingencies, the optimal solution given by (3) is first-best if and only if $\Pi \leq 1$. ■

When $\Pi \leq 1$, the above derived optimal solution is thus first-best. Moreover, there will exist no ex post incentives to deviate from the agreement, because losses from liberalization will never be incurred, while gains from liberalization will be made in one-quarter of trades (i.e. between any pair of country sectors that are both in the good state).

The categorization of sectors in the proof of the above proposition helps us interpret the somewhat surprising result in proposition 1 that the optimal adjustment phase length is gradually decreasing in Π, rather than jumping from infinity to zero at $\Pi = 1$. The underlying reason is that, under a clear-cut rule, there will be conflicting interests between the different categories of sectors when $\Pi > 1$. In the categories comprising corresponding sectors in different states in the two countries, it is globally optimal not to apply the safeguard, while in the category where corresponding sectors are simultaneously in the bad state, it is optimal to implement the safeguard. The optimal adjustment phase length obtained in proposition 1 thus strikes a balance between gains and losses made by these different categories of sectors under a safeguard with a predetermined time limit. As Π increases above one, the gain from not applying the safeguard in the two categories of corresponding sectors in different states increases. The optimal solution will thus increasingly be tilted in favor of the globally optimal solution for these two categories, i.e. $\widehat{\lambda}$ will fall, as Π rises. Eventually, when $\Pi \geq 2 + \frac{1}{1+2\delta}$, the gain for these two categories from having no safeguard outweighs the loss incurred for corresponding sectors simultaneously in the bad state and not allowed to deviate and hence, $\widehat{\lambda} = 0$ is optimal.

It immediately follows from the preceding discussion why the optimal solution is not first-best whenever $\Pi > 1$. The optimal solution can be seen as a compromise between the various categories of sectors. When $\Pi \in (1, 2 + \frac{1}{1+2\delta})$, it is the case that $\widehat{\lambda} \in (0, \infty)$ and thus, the safeguard will be applied if and only if at least one of any pair of corresponding sectors finds itself in state $\underline{\varepsilon}^{\lambda -}$. This is globally suboptimal, both for the two categories of corresponding sectors in different states, for which no safeguard at all would be optimal, and for corresponding sectors that are simultaneously in state $\underline{\varepsilon}^{\lambda +}$, in which case both countries would be better off scaling back liberalization, but neither of them is permitted to do so. When $\Pi \geq 2 + \frac{1}{1+2\delta}$ and hence $\widehat{\lambda} = 0$, global suboptimality only arises in the latter category, i.e. whenever corresponding sectors are simultaneously in state $\underline{\varepsilon}$.

The proof of the previous proposition suggests the first-best solution to be of bang-bang type. However, this is only true with some qualifications for $\Pi > 1$. The first-best solution when $\Pi > 1$ prescribes the safeguard to be used whenever corresponding sectors are simultaneously in the bad state, while not allowing for it to be applied when they are in different states. Hence, the safeguard should be applied if and only if and for as long as both countries would agree on this.

There are, however, strong practical and political reasons for assuming away the possibility of allowing for a safeguard to be applied, if and only if it has unanimous support. First, it will require the immediate withdrawal of the safeguard as soon as required by one country. More specifically, the implementability of the first-best solution will rest on liberalization being suspended only as long as both sectors are in the bad state; as soon as one country's sector switches to the good state, liberalization must be reintroduced, although that will make the country whose sector is still in the bad state worse off. This is complicated, because such an arrangement will amount to letting the duration of implementing a safeguard be determined by the state of the corresponding sector in the other country.[21] The duration of protection under a safeguard being dependent on the state abroad will make protection unpredictable. In the present context of entirely exogenously determined switches in states, this may be of minor importance. In reality, it could be politically difficult to agree upon such a rule. One of the benefits of having a clear-cut rule governing the length of the adjustment phase is that it makes the future predictable. A sector in need of protection will be able to relate its adjustment measures to the time limits

[21] This is in contrast to the principles governing the conduct of safeguards, which allow for their implementation in situations where an industry's situation is adversely affected (see subsection 2.1).

of the protection granted; letting the duration of protection depend upon factors beyond the control of this sector (stochastically determined in the context of this model) will create uncertainty.

Second, the first-best solution might be hard to enforce. Once liberalization had been suspended under mutual agreement (i.e when corresponding sectors are in state $\underline{\varepsilon}$), liberalization would have to be reintroduced as soon as one country's sector switched from the bad to the good state, although that would make the country whose sector would still be in the bad state, worse off. Ex post, there would thus be strong incentives for the country still in the bad state to continue applying the safeguard until this sector were to attain the good state. It would then be the case that once liberalization had been suspended, it could not be reintroduced before both sectors were once more in the good state.[22] In the present model, this may be less relevant, because an infinite number of sectors would make adhering to the first-best solution more easily enforceable.[23] Assuming a finite number of sectors, such consideration might, however, carry some weight.

Finally, it might be difficult to implement the first-best solution in a multi-country framework. The first-best solution might prescribe the safeguard to be applied only against countries with corresponding sectors also in the bad state, while countries with corresponding sectors in the good state would not be affected. This discriminatory use of the safeguard would amount to a breach against one of the fundamental principles of all trade agreements, the MFN clause.[24] In contrast, the derived optimal solution has the advantage of being easily reconciled with the MFN principle. Besides, the first-best solution might require coordination and supervision in a multi-country framework. A supranational monitoring or enforcement agency might be necessary, which might be costly and probably politically unfeasible. Apart from strong practicability constraints, it is highly unlikely that sovereign states

[22] A possible arrangement worth consideration is to allow for suspension of liberalization if commonly agreed upon and for liberalization not to be reintroduced until unanimously supported.

[23] The political economic context within a country would be of importance. If losers from liberalization (sectors in the bad state) are compensated by winners from liberalization (sectors in the good state), the first-best solution ought to be more easily adhered to, because, on aggregate, a country would gain from it.

[24] For rigorous theoretical support of the MFN principle, I refer to Bagwell and Staiger (1999). Horn and Mavroidis (2001) provide a survey of economic and legal aspects of the MFN clause.

would ever cede their control over trade policy to a supranational agency. The fact that countries have agreed to adopt rules governing the conduct of various provisions rather than letting trade policy be supranationally fine-tuned renders support to the underlying presumption that first-best solutions may not be feasible.

It can be concluded that implementing the first-best solution is problematic in reality when $\Pi > 1$, because it requires a rule prohibiting the use of the safeguard, except when this is mutually beneficial and, moreover, instantly removing the safeguard as soon as the better-off country no longer needs it. Whether this type of arrangement can be justified from an ex-ante point of view and hence, whether it is feasible, remains to be discussed. However, the preceding discussion suggests that a clear-cut rule, defining time limits on the implementation of safeguards, may be desirable. Instead of leaving it to the discretion of countries to apply the safeguard when it is globally optimal, a rule prespecifying the length of the adjustment phase under all contingencies may be preferable. Why a rule that is ex ante suboptimal in the present context is included, be it for practical reasons, or for political constraints, is beyond the scope of the analysis of this paper. Hence, the present framework does not provide a rationale for this assumption. The optimality conditions that are derived are thus subject to a rule, for which there exist strong political and practicability reasons not explicitly addressed in the present model.

5 Ex post Implications of a Time Limit

5.1 Current Payoffs under Optimality

When the optimal adjustment phase length is determined, the flow of future payoffs from the moment the agreement is implemented is maximized. As shown in the previous section, the optimal solution is having a safeguard without any upper time limit on its use if $\Pi \leq 1$, and including no safeguard provision at all if $\Pi \geq 2 + \frac{1}{1+2\delta}$. In these two cases, the average current payoff $\overline{w}$ is constant over time. If, however, $\Pi \in (1, 2 + \frac{1}{1+2\delta})$, the optimal solution prescribes a safeguard with an upper time limit on its use, i.e. $\widehat{\lambda} \in (0, \infty)$. In this case, both the average current payoff $\overline{w}$ and the value of the flow of average future payoffs $\overline{v}$ will change over time, once the agreement has been implemented. First, the effect on current payoffs will be assessed.

Lemma 2 *If* $\Pi \in (1, 2+\frac{1}{1+2\delta})$*, implementing the optimal adjustment phase length* $\widehat{\lambda}$ *results in* $\overline{w}$ *being constant for* $\tau \leq \widehat{\lambda}$*. As* τ *increases beyond* $\widehat{\lambda}$*,* $\overline{w}$ *initially increases, but eventually decreases, thereby converging to a value unambiguously larger than its value at the agreement's inception (*$\lim_{\tau\to\infty} \overline{w}(\tau, \widehat{\lambda}) > \overline{w}(0, \widehat{\lambda})$*).*

Proof. See the Appendix. ■

To get an intuitive understanding of this result, it is necessary to determine in which ways the average payoff $\overline{w}(\tau, \lambda)$ is affected over time. Let $\mu \equiv \mu(\underline{\varepsilon}^{\lambda-}, \tau)$, the share of sectors having been in state $\underline{\varepsilon}$ for a period shorter than λ. From above, we know that μ equals one-half for $\tau \leq \widehat{\lambda}$, decreases monotonically in τ for $\tau > \widehat{\lambda}$, and converges to $\frac{1-e^{-\widehat{\lambda}}}{2}$ as τ goes to infinity. There are three ways in which this change in μ over time has an impact on the average sector payoff. First, it unambiguously increases average payoffs in state $\overline{\varepsilon}$ for $\tau > \widehat{\lambda}$. Second, it unambiguously reduces average payoffs in state $\underline{\varepsilon}^{\widehat{\lambda}+}$ for $\tau > \widehat{\lambda}$. Third, it unambiguously depresses the average payoff as the share of sectors in state $\underline{\varepsilon}^{\widehat{\lambda}+}$ increases and the share of sectors in state $\underline{\varepsilon}^{\widehat{\lambda}-}$ decreases correspondingly. This negative effect due to the shift of sectors in state $\underline{\varepsilon}^{\widehat{\lambda}-}$ to state $\underline{\varepsilon}^{\widehat{\lambda}+}$ is exacerbated by the second effect that the average payoff in state $\underline{\varepsilon}^{\widehat{\lambda}+}$ decreases. To understand the impact of these different effects on $\overline{w}$ over time, $\overline{w}$ can be expressed as follows

$$\begin{aligned}\overline{w}(\tau, \lambda) &= \frac{1}{2}[\mu w_N(\overline{\varepsilon}) + (1-\mu) w_C(\overline{\varepsilon})] \\ &\quad + \mu w_N(\underline{\varepsilon}) + (\frac{1}{2} - \mu)[\mu w_N(\underline{\varepsilon}) + (1-\mu) w_C(\underline{\varepsilon})] \\ &= \frac{1}{2}\{w_C(\overline{\varepsilon}) - \mu\Pi[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\} \\ &\quad + \frac{1}{2} w_N(\underline{\varepsilon}) - (1-\mu)(\frac{1}{2} - \mu)[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})].\end{aligned}$$

The first effect enters via the first term, while the second and the third effects enter through the third term. While the first, positive effect is linear in μ, the combined impact of the last two, negative effects is quadratic in μ. It is easily established that the net marginal effect of a decrease in μ below one-half (i.e. at time $\tau = \widehat{\lambda}$) is strictly positive. Hence, $\overline{w}$ will increase beyond $\widehat{\lambda}$. As time passes and μ decreases further, the marginal impact on the third term becomes increasingly negative, since this term is quadratic in μ. Eventually, it will exceed the positive, constant marginal impact on the first term, and the net marginal effect will therefore become negative and $\overline{w}$ will start falling. Thus, the combined effect of bad-state sectors in need of extended protection becoming increasingly worse off on the one hand, and a shift of sectors from those protected by a safeguard to those that are not on the other hand, will eventually lead to a decline in the average sector payoff.

Using lemma 2, it is easily established that although both countries will eventually become increasingly worse off under the optimal adjustment phase length, it will nevertheless not be beneficial to breach the agreement and revert to the Nash equilibrium at any point in time.

Proposition 3 *An agreement with an optimal adjustment phase length is Nash-superior over the entire time horizon.*

Proof. Lemma 2 states that $\overline{w}$ is constant for $\tau \leq \widehat{\lambda}$ and initially increasing, before eventually decreasing as τ increases beyond $\widehat{\lambda}$ and, moreover, that $\lim_{\tau\to\infty} \overline{w}(\tau, \widehat{\lambda}) > \overline{w}(0, \widehat{\lambda})$. Hence, $\overline{w}(\tau, \widehat{\lambda}) \geq \overline{w}(0, \widehat{\lambda})$ for all $\tau \geq 0$. Since $\overline{w}(0, \widehat{\lambda}) > \frac{1}{2}w_N(\underline{\varepsilon}) + \frac{1}{2}w_N(\overline{\varepsilon})$, the value of discounted future flows of payoffs must be strictly larger under the optimal adjustment phase than in the absence of an agreement for any $\tau \geq 0$. Thus, at no point in time will there be any incentive to breach the agreement. ∎

While defecting from the agreement will never be worthwhile, it might nevertheless be the case that ex post, a different agreement might be preferable. To address that issue, the effect of limiting the duration of the safeguard on the discounted value of the flow of future payoffs over time must be assessed.

5.2 Flows of Future Payoffs under Optimality

As noted above, the discounted value of future flows of payoffs $\overline{v}$ is constant over time if $\Pi \leq 1$ and hence, $\widehat{\lambda} = 0$, or if $\Pi \geq 2+\frac{1}{1+2\delta}$ and hence, $\widehat{\lambda} = \infty$. For $\Pi \in (1, 2+\frac{1}{1+2\delta})$, however, it will be the case that $\widehat{\lambda} \in (0, \infty)$ and $\overline{v}$ varies over time. The lemma of the previous subsection, showing that under the optimal adjustment phase length $\widehat{\lambda}$ current payoffs are constant for $\tau \leq \widehat{\lambda}$, initially increase and then decrease as

τ increases beyond $\widehat{\lambda}$, suggests that the discounted value of future flows of payoff should initially increase and eventually fall. The following lemma demonstrates that the eventual decrease in $\overline{v}$ takes place after the agreement has been in place for a longer time than the adjustment phase length.

Lemma 3 *If* $\Pi \in (1, 2 + \frac{1}{1+2\delta})$*, implementing the optimal adjustment phase length* $\widehat{\lambda}$ *results in* $\overline{v}$ *increasing beyond* $\widehat{\lambda}$*, but eventually decreasing.*

Proof. See the Appendix. ■

From proposition 3, we know that the discounted value of future flows of payoffs will always be larger than in the absence of any agreement. The fact that $\overline{v}$ will eventually fall suggests that the agreed-upon time limit on the use of the safeguard might no longer be optimal ex post, thus leaving room for ex post adjustments. In what follows, it will be demonstrated that this is indeed the case.

5.3 Ex post Suboptimality

Since the discounted value of future flows of payoffs $\overline{v}$ is constant over time if no safeguard is included in the agreement ($\widehat{\lambda} = 0$), it immediately follows that if $\widehat{\lambda} = 0$ is optimal at the agreement's inception, it will be so over the infinite time horizon. The same applies to the case when a safeguard with no time limit ($\widehat{\lambda} = \infty$) is implemented. For $\widehat{\lambda} \in (0, \infty)$, however, $\overline{v}$ will vary over time and it may be the case that $\widehat{\lambda}$ is no longer optimal, once the agreement has been implemented.

By introspection of (2), it is easily seen that the optimal solution is independent of τ for $\tau \leq \lambda$. Hence, the optimal solution is also ex post optimal as long as $\tau \leq \lambda$. For $\tau > \lambda$, this is no longer the case, however, as demonstrated by the following proposition.

Proposition 4 *If* $\Pi \in (1, 2+\frac{1}{1+2\delta})$*, the ex post optimal solution for* λ *will eventually be higher than the agreed-upon solution* $\widehat{\lambda}$*. As* τ *increases, so will the ex post optimal solution for* λ*.*

Proof. See the Appendix. ■

Eventually, countries would be better off, if a higher λ had been chosen at the agreement's inception.[25] The intuition behind this result is that when the agreement is implemented, insufficient weight is attributed to the category of sectors that have been in the bad state for longer than the length of the adjustment phase and hence, are in need of further protection. Since the share of this category of sectors is zero in the initial phase of the agreement, i.e. as long as $\tau \leq \widehat{\lambda}$, and then increases, the weight it is given is too small ex ante.

It is important to emphasize that the ex post optimal adjustment phase length is optimal in the sense that it would yield the highest value of discounted flows of future payoffs at a specific point in time. A government will thus ex post perceive that a higher value of discounted future payoff flows could have been attained if a different adjustment phase length had been chosen. Hence, the ex ante suboptimality associated with including a clear-cut rule prescribing an upper time limit on the use of the safeguard will also lead to ex post suboptimality.

The agreed-upon time limit for the use of the safeguard will thus eventually be perceived to be too short. Or, in other words, dissatisfaction will grow over time. The demands for increasing the time limit might increase, which implies that there may be some room for ex post renegotiation to modify the adjustment phase length. In the present model, where countries are equally well off due to an infinite number of sectors, getting unanimous support for any ex post readjustments would be no problem. If the number of sectors were finite, however, countries might not be equally well off and hence, agreement on ex post readjustments might be harder to achieve.

It is, however, important to emphasize that the optimal solution derived in the previous section implicitly rests on the assumption that ex post readjustments are not possible, or in other words, that governments can commit to the rules of the agreement. If the time limit could be modified ex post, governments should correctly anticipate such adjustments ex ante, which would feed back into the determination

[25] In the proof of proposition 4, it is shown that the ex post optimal time limit will actually be strictly smaller than $\widehat{\lambda}$ as τ increases beyond $\widehat{\lambda}$. The underlying reason seems to be that there is a kink in the curve of $\overline{v}$ at $\tau = \widehat{\lambda}$. While the observation that countries will initially perceive $\widehat{\lambda}$ to be too large as τ increases beyond $\widehat{\lambda}$ is interesting, it will not be addressed any further. From a practical point of view, there are likely to be obstacles to ex post lower the adjustment phase length. Anyway, the ex post optimal time limit will eventually be increasingly larger than $\widehat{\lambda}$.

of the time limit being implemented at the inception of the agreement. In fact, if readjustments were possible, an optimal solution would prescribe adjustment taking place such that optimality would be satisfied at any point in time. Such solutions are obviously an interesting topic of future research.

More generally, however, the fact that ex post subpotimality will arise in the presence of a finite, time limit on the use of safeguards implies that the obtained solution is not time consistent. Ex post, i.e. for $\tau > \widehat{\lambda}$, the originally adopted solution will no longer yield the highest possible value of flows of discounted future payoffs and hence, a change in the time limit would be beneficial.[26]

Whether an ex post adjustment of the time limit on the use of the safeguard is feasible remains to be investigated. An interesting situation would arise if, at some point in time, the discounted future flow of payoffs were to become lower than at the moment when the agreement is implemented. In this case, it would be optimal to restart the agreement. The following proposition states that this is always the case when the optimal solution is applied.

Lemma 4 *If* $\Pi \in (1, 2 + \frac{1}{1+2\delta})$*, implementing the optimal adjustment phase length* $\widehat{\lambda}$ *results in* $\overline{v}(\tau, \widehat{\lambda}) < \overline{v}(0, \widehat{\lambda})$ *for sufficiently large* τ*.*

Proof. See the Appendix. ■

This proposition has an important implication. It will be the case that the discounted value of future payoff flows is eventually lower than its initial value (at the agreement's inception). Once this is the case, reimplementing the agreement will thus be beneficial. However, as pointed out above, if it could be anticipated that the agreement be restarted, this would have to be taken into account in the solution taken at the inception of the agreement. Moreover, restarting the agreement might be problematic, because it would make current payoffs lower, although the discounted value of the future flows of payoffs would become unambiguously larger. Furthermore, a reimplementation of the agreement would create winners and losers, the beneficiaries being sectors in state $\underline{\varepsilon}^{\lambda+}$, and to a lesser extent sectors in state

[26] The present time inconsistency seems to originate in having a rigid rule prescrining a fixed time limit to being exempted from a commitment, thereby suggesting a similarity to the time inconsistency identified in Kydland and Prescott (1977).

$\underline{\varepsilon}^{\lambda-}$, and the maleficiaries being sectors in state $\overline{\varepsilon}$, something that could make a restart politically controversial.

The ex post suboptimality identified here rests on the assumption of an infinite number of sectors. If the number of sectors were instead finite, aggregate payoffs would differ from the ex ante expected value and hence, ex post dissatisfaction would not necessarily arise. However, since the derived expressions for $\overline{v}$ would still apply as expected values, it could ex ante be expected that ex post suboptimality might arise.

The result of ex post suboptimality of a time limit on being exempted from a cooperative arrangement renders an interesting interpretation of the ongoing controversy about the growth and stability pact in the euro area, an agreement under which participating countries are obliged to adhere to certain prespecified rules. In particular, the pact prescribes budget deficits not to exceed 3% of GDP in more than three consecutive years. At the time when the pact was negotiated, the prospect of a country finding itself in the situation of needing to run a deficit exceeding 3% of GDP in a fourth consecutive year was in the distant future and therefore, low weight was attributed to this possibility. It can be argued that while the budget deficit rule seemed optimal when the pact was signed, it is increasingly being regarded as too rigid. In fact, during recent years, several participating countries have found themselves in the position of risking to run a deficit exceeding 3% of GDP for more than three years in a row. Not surprisingly, voices have been raised for relaxing this rule, or at least interpreting it more generously.

To conclude the analysis of the case with two countries, each with an infinite number of sectors subject to stochastic variation in states, a safeguard with a time limit is associated with both ex ante and ex post inefficiencies. Next, the model will be modified such that asymmetry across countries arises, and it will be shown that similar results can be obtained.

6 The Asymmetric Case

In contrast to the previously investigated case with both countries being subject to stochastic shifts between states, asymmetry will now be introduced. The two countries will henceforth be referred to as the home and the foreign (distinguished by an asterisk) country. For simplicity, it will be assumed that both countries are symmetric in all sectors but one, and that the symmetric sectors are not subject to stochastic switches between states and hence, are stable over time. In the remaining sector, it will, however, be the case that the home country is exposed to switches between states, determined by the same stochastic process as above, while the corresponding sector in the foreign country is stable over time. In what follows, this sector will be referred to as the asymmetric sector.

This way of introducing asymmetry could, for example, be justified in a framework where the home country is a developing country, the asymmetric sector being an infant industry, while the foreign country is a developed country with no need for protection in that sector. The case with a large number of infant industry sectors, as well as the case with both countries having infant industries, albeit in different sectors, will be discussed at the end of the section.

An important analytical implication of only one sector in one country being exposed to stochastic shifts between states is that it is no longer possible to determine the average discounted value of future flows of payoffs for any moment in time, as in the previous symmetric case with an infinite number of sectors. Instead, expected values of future flows of payoffs in the asymmetric sector will be calculated, and these expected values will be state-dependent and differ between countries. In fact, if the assumption of only one (or any finite number) sector had been made in the previous sections, expected state-contingent values of future flows of payoffs would have been obtained. Conversely, an infinite number of asymmetric sectors would make the average discounted value of future flows of payoffs over time independent of sector states. The approach taken in this section is thus more relevant for a low number of sectors (e.g. one), when actual outcomes may differ from expected outcomes, and can hence be seen as complementary to the previous approach, which applies to a large (possibly infinite) number of sectors, where actual outcomes are very likely to be close to the expected outcomes.

It suffices to focus on the sector subject to switches between states, because all other sector payoffs are constant over time. In what follows, only current payoffs and discounted values of future flows of payoffs in the asymmetric sector will therefore be considered. While assumptions 1 and 3 still apply, assumption 2 is modified as follows.

Assumption 2'

$$w_C(\overline{\varepsilon}) > w_N(\overline{\varepsilon})$$

$$w_C(\underline{\varepsilon}) < w_N(\underline{\varepsilon})$$

$$w_C^* > w_N^*.$$

Thus, it is the case that the foreign country is unambiguously better off under liberalization, while the home country, as before, is better off under liberalization if and only if it finds itself in the good state. Depending on in which state it currently is, the home country's expected value of the discounted flow of payoffs in the absence of any trade agreement v_N is given by (see the Appendix for derivation)

$$\begin{aligned} v_N(\underline{\varepsilon}) &= \frac{1}{\delta(2+\delta)}[(1+\delta)w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})] \\ v_N(\overline{\varepsilon}) &= \frac{1}{\delta(2+\delta)}[w_N(\underline{\varepsilon}) + (1+\delta)w_N(\overline{\varepsilon})]. \end{aligned}$$

As expected, the continuation value when the home country is in state $\overline{\varepsilon}$ is strictly larger than when the home country is in state $\underline{\varepsilon}$. The foreign country's value of discounted flows of future payoffs in the absence of an agreement is, as above, given by

$$v_N^* = \int_0^\infty w_N^* e^{-\delta t} dt = \frac{1}{\delta} w_N^*.$$

Next, trade liberalization will be introduced.

6.1 Introducing a Trade Agreement

Since the home country's asymmetric sector may actually be worse off under trade liberalization, its government will push for the inclusion of a safeguard without any time limit, while the foreign country will be in favor of no safeguard at all, because it always gains from liberalization. Different from the previously investigated case with stochastic switches between states in both countries, the two countries will now have opposite interests regarding the length of the adjustment phase. Any agreed-upon time limit on the use of the safeguard must therefore be negotiated through bargaining.

Before analyzing the bargaining process, it is necessary to calculate the continuation values in the presence of a safeguard with a time limit of λ. The home country's value of the discounted flow of payoffs under a trade agreement v_C will depend on in which state it is. Being in state $\underline{\varepsilon}$, it will also depend on whether the home country uses the safeguard and how long the safeguard has been applied. Let T be the time the home country has been in state $\underline{\varepsilon}$. The continuation values are then given by (see the Appendix for derivation)

$$v_C(\underline{\varepsilon}, T) = \begin{cases} \frac{1+\delta}{\delta(2+\delta)} w_N(\underline{\varepsilon}) + \frac{1}{\delta(2+\delta)} w_C(\overline{\varepsilon}) & \text{for } T \in [0, \lambda] \\ -\frac{\delta(2+\delta)e^{-(1+\delta)(\lambda - T)} + e^{-(1+\delta)\lambda}}{\delta(1+\delta)(2+\delta)} [w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] & \\ & \\ \frac{1+\delta}{\delta(2+\delta)} w_C(\underline{\varepsilon}) + \frac{1}{\delta(2+\delta)} w_C(\overline{\varepsilon}) & \\ +\frac{1-e^{-(1+\delta)\lambda}}{\delta(1+\delta)(2+\delta)} [w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] & \text{for } T > \lambda \end{cases}$$

$$v_C(\overline{\varepsilon}) \quad = \frac{1}{\delta(2+\delta)} w_C(\underline{\varepsilon}) + \frac{1+\delta}{\delta(2+\delta)} w_C(\overline{\varepsilon}) + \frac{1-e^{-(1+\delta)\lambda}}{\delta(2+\delta)} [w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\}.$$

The foreign country's continuation values are given by (see the Appendix for derivation)

$$v_C^*(\underline{\varepsilon}, T) = \begin{cases} \frac{1}{\delta(2+\delta)} w_C^* + \frac{1+\delta}{\delta(2+\delta)} w_N^* & \text{for } T \in [0, \lambda] \\ +\frac{\delta(2+\delta)e^{-(1+\delta)(\lambda - T)} + e^{-(1+\delta)\lambda}}{\delta(1+\delta)(2+\delta)} (w_C^* - w_N^*) & \\ & \\ \frac{1}{\delta} w_C^* - \frac{1-e^{-(1+\delta)\lambda}}{\delta(1+\delta)(2+\delta)} (w_C^* - w_N^*) & \text{for } T > \lambda \end{cases}$$

$$v_C^*(\overline{\varepsilon}) \quad = \frac{1}{\delta} w_C^* - \frac{1-e^{-(1+\delta)\lambda}}{\delta(2+\delta)} (w_C^* - w_N^*).$$

When no safeguard is allowed ($\lambda = 0$), $v_C(\underline{\varepsilon}, T) = \frac{(1+\delta)w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon})}{\delta(2+\delta)}$ and $v_C(\overline{\varepsilon}) = \frac{w_C(\underline{\varepsilon}) + (1+\delta)w_C(\overline{\varepsilon})}{\delta(2+\delta)}$, while $v_C^*(\underline{\varepsilon}, T) = \frac{w_C^*}{\delta}$ and $v_C^*(\overline{\varepsilon}) = \frac{w_C^*}{\delta}$. In case there is no upper time limit on the use of the safeguard ($\lambda = \infty$), $v_C(\underline{\varepsilon}, T) = \frac{(1+\delta)w_N(\underline{\varepsilon}) + w_C(\overline{\varepsilon})}{\delta(2+\delta)}$ and $v_C(\overline{\varepsilon}) = \frac{w_N(\underline{\varepsilon}) + (1+\delta)w_C(\overline{\varepsilon})}{\delta(2+\delta)}$, while $v_C^*(\underline{\varepsilon}, T) = \frac{w_C^* + (1+\delta)w_N^*}{\delta(2+\delta)}$ and $v_C^*(\overline{\varepsilon}) = \frac{(1+\delta)w_C^* + w_N^*}{\delta(2+\delta)}$.

When a safeguard of strictly positive, but finite length is implemented ($\lambda \in (0, \infty)$), $v_C(\underline{\varepsilon}, T)$ falls monotonously in T, as long as the home country is allowed to apply the safeguard; the closer it gets to the time limit of the adjustment phase, the lower its continuation value will be. Once the duration of being in the bad state exceeds the adjustment phase length, the continuation value is constant. By the same token, $v_C^*(\underline{\varepsilon}, T)$ increases monotonously in T as long as the home country is allowed to apply the safeguard and remains constant thereafter.

Since a more generous adjustment phase length unambiguously benefits the home country, while adversely affecting the foreign country, it is straightforward that the home country's continuation values increase in λ, while the foreign country's continuation values decrease in λ.

For a cooperative agreement to be sustainable, it is necessary that switching to the Nash equilibrium when not permitted under the agreement does not increase the expected payoffs. For the foreign country, this can never be the case, since it is always better off under cooperation than under the Nash equilibrium. The home country might, however, prefer infinite Nash reversion (assuming grim-trigger strategies against deviations) when having been in the bad state for longer than the adjustment phase length, depending on the discounting rate. If breaching the agreement by extending protective measures beyond the time limit of the safeguard can be punished by reversion of liberalization in other sectors, agreement-conform behavior can be induced. Thus, the number of sectors not subject to stochastic fluctuations will be assumed to be sufficiently large to make sustaining the agreement worthwhile and hence, no participation constraint on the adjustment phase length need to be considered.[27]

6.2 The Nash Bargaining Solution

It will be assumed that negotiations between the home and foreign countries are pursued sector-wise, and that the two countries engage in Nash bargaining over the degree of liberalization and the maximally permitted adjustment phase length λ in the asymmetric sector. Let v_C^e and v_C^{*e} be the ex ante expected continuation values

[27] Alternatively, the discount rate could be assumed to be sufficiently low to make infinite Nash reversion inferior to complying with the agreement.

in the asymmetric sector at the agreement's inception. Assuming an ex ante equal likelihood of $\underline{\varepsilon}$ and $\overline{\varepsilon}$ occurring when the agreement is implemented, the following expressions are obtained

$$v_C^e = \frac{1}{2\delta}\{w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) + (1 - e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\}$$
$$v_C^{*e} = \frac{1}{2\delta}\{2w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)\}.$$

Let λ_{NBS} be the optimal solution under Nash bargaining, henceforth referred to as the Nash bargaining solution (NBS).

Proposition 5 *The NBS for the adjustment phase length decreases in the ratio between the home country's gains and losses from liberalization. More specifically,*

$$\lambda_{NBS} = \begin{cases} \infty & \text{if } \Pi \leq 1 \\ \frac{\ln 2 - \ln(\Pi - 1)}{1+\delta} & \text{if } \Pi \in (1,3) \\ 0 & \text{if } \Pi \geq 3 \end{cases} . \tag{4}$$

Proof. See the Appendix. ■

Somewhat surprisingly, the foreign country's gain from cooperation, given by $w_C^* - w_N^*$, does not have any impact on λ_{NBS}. The NBS will solely depend on the ratio between the gain in state $\overline{\varepsilon}$ and the loss in state $\underline{\varepsilon}$ by the home country, given a certain degree of liberalization. Hence, if the gains from cooperation are smaller than the losses for the home country, its demand for unlimited protection in bad times is non-negotiable. If the gains from cooperation are much larger (by a factor larger than three) than the losses for the home country, the foreign country's demand for no protection at all for the home country in bad times will be non-negotiable. It is only when the gains from cooperation are somewhat larger (by a factor between one and three) than the losses for the home country that negotiations will produce an adjustment phase that is strictly positive but also finite.

This result is similar, although not equivalent, to that obtained in the previous case with stochastic switches between states in both countries. Proposition 1 stated that the optimal length of the adjustment phase $\widehat{\lambda}$ decreases monotonously in Π.

Moreover, it was shown in lemma 1 that $\widehat{\lambda}$ decreases in the discounting rate δ. The same is true for the NBS with some qualifications, as shown in the following lemma.

Lemma 5 *For $\Pi \in (1, 3)$, the optimal solution under Nash bargaining λ_{NBS} unambiguously decreases in the discount rate. For $\Pi \leq 1$ or $\Pi \geq 3$, however, the discount rate has no impact on λ_{NBS}.*

Proof. Follows immediately by introspection of (4). ■

To intuitively understand this result, consider the effects of a change in δ on v_C^e and v_C^{*e}. Since v_C^e unambiguously increases and v_C^{*e} unambiguously decreases in δ, the NBS will lie closer to what is considered optimal by the foreign country, i.e. λ_{NBS} will fall, if δ increases.[28] If $\lambda_{NBS} = 0$ or $\lambda_{NBS} = \infty$ to start with, then a change in the discount rate has no effect on v_C^e and v_C^{*e} and hence, has no effect on the outcome of Nash bargaining either.

6.3 The Global Optimum

Next, the NBS will be contrasted to the globally optimal solution, denoted by λ_{opt} and obtained by maximizing home and foreign payoffs in the asymmetric sector.

Proposition 6 *The globally optimal solution is given by*

$$\lambda_{opt} = \begin{cases} \infty & \text{if } w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) > w_C^* - w_N^* \\ [0, \infty] & \text{if } w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) = w_C^* - w_N^* \\ 0 & \text{if } w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) < w_C^* - w_N^* \end{cases} .$$

Proof. The globally optimal solution is obtained by maximizing $v_C^e + v_C^{*e}$.

$$v_C^e + v_C^{*e} = \frac{1}{2\delta}\{w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) + 2w_C^* + (1 - e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) - (w_C^* - w_N^*)]\}$$

The proposition immediately follows by introspection of this equation. ■

[28] Note that although λ_{NBS} decreases in the discount rate when $\Pi \in (1, 3)$, it will never become zero, in contrast to $\widehat{\lambda}$, which may eventually fall to zero.

The result is straightforward. If the home country's loss while applying cooperative tariffs in state $\underline{\varepsilon}$ is smaller (larger) than the foreign country's gain from cooperation, the adjustment phase should be as short (long) as possible. Hence, a bang-bang solution is obtained. The globally optimal solution is thus equivalent to the first-best solution in the symmetric case where all sectors are subject to switches between states, except that no coordination on the use of the safeguard is necessary.

The following lemma contrasts λ_{opt} and λ_{NBS}.

Lemma 6 *When* $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) < w_C^* - w_N^*$ *then the following is true:*
(i) If $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) \leq \frac{1}{3}[w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})]$, *then* $\lambda_{opt} = \lambda_{NBS} = 0$.
(ii) If $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) > \frac{1}{3}[w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})]$, *then* $0 = \lambda_{opt} < \lambda_{NBS} \leq \infty$.
When $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) > w_C^* - w_N^*$ *then the following is true:*
(i) If $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) < w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})$, *then* $0 \leq \lambda_{NBS} < \lambda_{opt} = \infty$.
(ii) If $w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) \geq w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})$, *then* $\lambda_{NBS} = \lambda_{opt} = \infty$.

Hence, the difference between the Nash bargaining outcome and what is globally optimal can be huge. The global inefficiencies that may arise under the NBS could justify some modification of the safeguard instrument. If $w_C^* - w_N^* > w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) > \frac{1}{3}[w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})]$ (implying $\Pi < 3$), then $\lambda_{NBS} > \lambda_{opt} = 0$. Since in this case having no safeguard is globally optimal, but the home country will be worse off under liberalization if its asymmetric sector finds itself in the bad state, efficiency could be enhanced if, instead of scaling back liberalization by invoking the safeguard, the home country would receive some compensation whenever it loses from liberalization.

Likewise, if $w_C^* - w_N^* < w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon}) < w_C(\overline{\varepsilon}) - w_N(\overline{\varepsilon})$ (implying $\Pi > 1$), and hence $\lambda_{NBS} < \lambda_{opt} = \infty$, global efficiency could be enhanced by allowing for the unlimited use of the safeguard, while compensating the foreign country for its foregone gains from liberalization, instead of having a time limit on the use of the safeguard. In fact, this case would be equivalent to attaching a cost to the use of the safeguard. In Herzing (2005b), the cost for applying the safeguard is necessary to restrain its use, whereby efficiency is also enhanced. Here, allowing for the unlimited application of the safeguard while appropriately compensating the trading partner would be Pareto superior to having a time limit on the use of the safeguard without any compensation.

In case the switching of states were entirely random but would also depend on the efforts undertaken, these efficiency-enhancing solutions could be susceptible to incentive compatibility problems, however. Arguably, the latter arrangement, with transfers from the home to the foreign country, would be easier to implement than the former, with transfers from the foreign to the home country, because it would not involve such problems; the home country would have no interest in transferring compensation to the foreign country for a longer period than necessary.

While increasing the number of sectors where the home country is subject to switches in states will not qualitatively change the NBS, it will affect the average expected flow of future sector payoffs in both countries. Letting the number of asymmetric sectors go to infinity, the average expected flows of future sector payoffs will decrease in the home country and increase in the foreign country over time, because the use of the safeguard will decrease over time, just as in the symmetric case. Hence, the ex post NBS for the adjustment phase will increase over time. Thus, dissatisfaction regarding the adjustment phase length will increase in the home country, where it will be perceived as being too short, while the foreign country will become increasingly well off. Ex post readjustments will therefore be hard to agree upon. This is in contrast to the symmetric case, where dissatisfaction will increase in both countries.

7 Conclusion

This paper has attempted to shed light on the determination of time limits for how long countries should be allowed to withdraw commitments made under a trade agreement. A symmetric two-country model, where sectors stochastically switch between two states, has been applied. While there are gains to be made from liberalization in the good state, losses will be incurred under liberalization in the bad state. It may therefore be desirable to agree on a rule, prescribing how long protection may be granted to sectors in bad states. It has been shown that the optimal time limit on protection will depend on the ratio between the gain from liberalization under the good state and the loss from liberalization under the bad state. For low ratios, no upper limit on protection is optimal and for high ratios, allowing for no protection at all is optimal. For intermediate ratios, however, a strictly positive and finite time limit is optimal. The optimal time limit thus serves to balance the interests of winners and losers in liberalization across countries.

Whenever the ex ante agreed-upon time limit is strictly positive and finite, it will eventually be perceived as being too low to an increasing extent. Thus, countries will ex post find the agreed upon adjustment phase length suboptimal. Hence, the dissatisfaction with the agreement will grow over time.

In a modified version of the model, asymmetry is introduced such that stochastical switches between states will occur in only one country. Assuming Nash bargaining between the two countries, a similar solution as in the case of symmetry is obtained. It is shown that the optimal time limit on protection will depend on the ratio between the gain from liberalization under the good state and the loss from liberalization under the bad state in the country subject to stochastic switches between states. As in the symmetric case, no upper limit on protection is optimal for low ratios, allowing for no protection at all is optimal for high ratios, and a strictly positive and finite time limit is optimal for intermediate ratios.

The two cases that have been examined can be seen as complementary. While the case of symmetry can be regarded as relevant for similar countries, i.e. developed countries, the case of asymmetry may be seen as more relevant for dissimilar countries, i.e. one developed and one developing country. Interestingly, the results obtained are similar.

The present analysis has entirely focused on the optimal length of a safeguard allowing for the withdrawal of liberalization commitments. No link to the actual extent of these commitments was established. An important topic for further research is to relate the optimal adjustment phase length to the degree of liberalization agreed-upon in trade negotiations. For that purpose, it is necessary to apply a more specific model, from which the interrelation between liberalization and the associated potential gains and losses can be derived. If, for instance, the ratio between potential gains and losses from liberalization will decrease in the degree of liberalization, the optimal adjustment phase length will increase in the degree of liberalization. Such an outcome would correspond well to the common perception that, by increasing the exposure to world markets, more liberalization ought to be combined with more flexibility, in particular increasing possibilities to scale back liberalization, for example through more generous time limits on safeguards.

8 References

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

Bagwell, Kyle and Staiger, Robert W. (1999), *An Economic Theory of GATT*, American Economic Review, 89(1), 215-248

Bagwell, Kyle and Staiger, Robert W. (2002), *The Economics of the World Trading System*, The MIT Press, Cambridge, Massachusetts

Ethier, Wilfried J. (2001), *Theoretical Problems in Negotiating Trade Liberalization*, European Journal of Political Economy, 17, 209-232

Ethier, Wilfried J. (2002), *Escape and Entry Mechanisms in the Multilateral Trade System*, PIER Working Paper 02-009, Department of Economics, University of Pennsylvania

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

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

Herzing, Mathias (2005a), *Can Self-Destructive Trade Agreements Be Optimal?*, mimeo, Institute for International Economic Studies, Stockholm University, Sweden

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

Hochman, Gal (2004), *The Interdependence of Safeguard Measures and Tariffs*, mimeo, William Davidson Faculty of Industrial Engineering and Management, Technion, Israel

Horn, Henrik and Mavroidis, Petros C. (2001), *Economic and Legal Aspects of the Most-Favored-Nation Clause*, European Journal of Political Economy, 17, 233-279

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

Kydland, Finn E. and Prescott, Edward C. (1977), *Rules Rather than Discretion: The Inconsistency of Optimal Plans*, Journal of Political Economy, 85(3), 473-491

Martin, Alberto and Vergote, Wouter (2004), *A Case for Strategic Antidumping*, mimeo, Columbia University, New York

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

Rosendorff, B. Peter and Milner, Helen (2001), *The Optimal Design of International Trade Institutions: Uncertainty and Escape*, International Organization, 55(4), 829-857

World Trade Organization (1994a), *The General Agreement on Tariffs and Trade*, Geneva, Switzerland

World Trade Organization (1994b), *Agreement on Safeguards*, Geneva, Switzerland

9 Appendix

9.1 Derivation of $\overline{v}(\tau,\lambda)$

The expected discounted flow of future payoffs for $\tau < \lambda$ is given by

$$\begin{aligned}
\overline{v}(\tau,\lambda) &= \int_0^\infty \overline{w}(\tau+t)e^{-\delta t}dt \\
&= \int_0^{\lambda-\tau} \frac{1}{4}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})]e^{-\delta t}dt \\
&\quad + \int_{\lambda-\tau}^\infty \left\{ \begin{array}{l} \frac{1}{2}w_C(\underline{\varepsilon}) + \frac{1}{2}w_C(\overline{\varepsilon}) \\ +\frac{1}{4}\frac{1-e^{-\lambda}}{1-e^{-\tau-t}}[3 - \frac{1-e^{-\lambda}}{1-e^{-\tau-t}} - \Pi][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \end{array} \right\} e^{-\delta t}dt \\
&= \frac{1-e^{-\delta(\lambda-\tau)}}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\delta(\lambda-\tau)}}{2\delta}[w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon})] \\
&\quad + \frac{1-e^{-\lambda}}{4}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \int_{\lambda-\tau}^\infty [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}]e^{-\delta t}dt.
\end{aligned}$$

The integral can be transformed as follows.

$$\begin{aligned}
&\int_{\lambda-\tau}^\infty [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}]e^{-\delta t}dt \\
&= [3-\Pi-(1-e^{-\lambda})]\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt - (1-e^{-\lambda})\int_{\lambda-\tau}^\infty \frac{e^{-\tau-t}e^{-\delta t}}{(1-e^{-\tau-t})^2}dt \\
&= [3-\Pi-(1-e^{-\lambda})]\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt \\
&- (1-e^{-\lambda})\{[\frac{-e^{-\delta t}}{1-e^{-\tau-t}}]_{\lambda-\tau}^\infty - \delta\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt]\} \qquad (*) \\
&= [3-\Pi-(1-\delta)(1-e^{-\lambda})]\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt - e^{-\delta(\lambda-\tau)}
\end{aligned}$$

The remaining integral is calculated as follows.

$$\begin{aligned}
\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt &= \int_{\lambda-\tau}^\infty e^{-\delta t}\sum_{j=0}^\infty e^{-j(\tau+t)}dt = \sum_{j=0}^\infty \int_{\lambda-\tau}^\infty e^{-\delta t-j(\tau+t)}dt \\
&= \sum_{j=0}^\infty [-\frac{e^{-\delta t-j(\tau+t)}}{\delta+j}]_{\lambda-\tau}^\infty = \frac{e^{-\delta(\lambda-\tau)}}{\delta}\sum_{j=0}^\infty \frac{\delta e^{-j\lambda}}{\delta+j} \qquad (\#)
\end{aligned}$$

$$= \frac{e^{-\delta(\lambda-\tau)}}{\delta} \frac{1}{1-\frac{\delta e^{-\lambda}}{\delta+1}} = \frac{1+\delta}{\delta} \frac{e^{-\delta(\lambda-\tau)}}{1+\delta(1-e^{-\lambda})}$$

Hence,

$$\begin{aligned}
&\int_{\lambda-\tau}^{\infty} [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}] e^{-\delta t} dt \\
&= \frac{e^{-\delta(\lambda-\tau)}}{\delta[1+\delta(1-e^{-\lambda})]} \{(1+\delta)[3-\Pi-(1-\delta)(1-e^{-\lambda})] - \delta[1+\delta(1-e^{-\lambda})]\} \\
&= \frac{e^{-\delta(\lambda-\tau)}}{\delta[1+\delta(1-e^{-\lambda})]} [(1+\delta)(2-\Pi)+e^{-\lambda}].
\end{aligned}$$

Thus the expected discounted flow of future payoffs for $\tau < \lambda$ is given by

$$\begin{aligned}
\overline{v}(\tau,\lambda) &= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\delta(\lambda-\tau)}}{4\delta}(\Pi-2)[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\
&\quad + \frac{e^{-\delta(\lambda-\tau)}}{4\delta[1+\delta(1-e^{-\lambda})]}(1-e^{-\lambda})[(1+\delta)(2-\Pi)+e^{-\lambda}][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\
&= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] \\
&+ \frac{e^{-\delta(\lambda-\tau)}}{4\delta[1+\delta(1-e^{-\lambda})]} \left\{ \begin{array}{c} (\Pi-2)[1+\delta(1-e^{-\tau})] \\ +(1+\delta)(1-e^{-\lambda})(2-\Pi) \\ +e^{-\lambda}(1-e^{-\lambda}) \end{array} \right\} [w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\
&= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] \\
&\quad + \frac{[\Pi-1-e^{-\lambda}]e^{-\delta(\lambda-\tau)}e^{-\lambda}}{4\delta[1+\delta(1-e^{-\lambda})]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})].
\end{aligned}$$

The expected discounted flow of future payoffs for $\tau \geq \lambda$ is given by

$$\begin{aligned}
\overline{v}(\tau,\lambda) &= \int_0^{\infty} \overline{w}(\tau+t) e^{-\delta t} dt \\
&= \int_0^{\infty} \left\{ \begin{array}{c} \frac{1}{2}w_C(\underline{\varepsilon}) + \frac{1}{2}w_C(\overline{\varepsilon}) \\ +\frac{1}{4}\frac{1-e^{-\lambda}}{1-e^{-\tau-t}}[3 - \frac{1-e^{-\lambda}}{1-e^{-\tau-t}} - \Pi][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \end{array} \right\} e^{-\delta t} dt \\
&= \frac{1}{2\delta}[w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon})] \\
&\quad + \frac{1-e^{-\lambda}}{4}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \int_0^{\infty} [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}] e^{-\delta t} dt.
\end{aligned}$$

The integral can be transformed as follows.

$$\int_0^\infty [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}]e^{-\delta t}dt$$

$$\begin{aligned}
&\overset{(*)}{=} [3-\Pi-(1-e^{-\lambda})]\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt \\
&\quad -(1-e^{-\lambda})\{[\frac{-e^{-\delta t}}{1-e^{-\tau-t}}]_0^\infty - \delta\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt]\} \\
&= [3-\Pi-(1-\delta)(1-e^{-\lambda})]\int_{\lambda-\tau}^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt - \frac{1-e^{-\lambda}}{1-e^{-\tau}}
\end{aligned}$$

The remaining integral is calculated as follows.

$$\begin{aligned}
\int_0^\infty \frac{e^{-\delta t}}{1-e^{-\tau-t}}dt &\overset{(\#)}{=} \sum_{j=0}^\infty [-\frac{e^{-\delta t - j(\tau+t)}}{\delta+j}]_0^\infty = \frac{1}{\delta}\sum_{j=0}^\infty \frac{\delta e^{-j\tau}}{\delta+j} \\
&= \frac{1}{\delta}\frac{1}{1-\frac{\delta e^{-\tau}}{\delta+1}} = \frac{1+\delta}{\delta}\frac{1}{1+\delta(1-e^{-\tau})}
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int_0^\infty [\frac{3-\Pi}{1-e^{-\tau-t}} - \frac{1-e^{-\lambda}}{(1-e^{-\tau-t})^2}]e^{-\delta t}dt \\
&= [3-\Pi-(1-\delta)(1-e^{-\lambda})]\frac{1+\delta}{\delta}\frac{1}{1+\delta(1-e^{-\tau})} - \frac{1-e^{-\lambda}}{1-e^{-\tau}} \\
&= \frac{1}{\delta[1+\delta(1-e^{-\tau})]}\left\{ \begin{array}{c} (1+\delta)[3-\Pi-(1-\delta)(1-e^{-\lambda})] \\ -\delta[1+\delta(1-e^{-\tau})]\frac{1-e^{-\lambda}}{1-e^{-\tau}} \end{array} \right\} \\
&= \frac{1}{\delta[1+\delta(1-e^{-\tau})]}\{(1+\delta)(2-\Pi)+e^{-\lambda}-\delta\frac{1-e^{-\lambda}}{1-e^{-\tau}}\}.
\end{aligned}$$

Thus the expected discounted flow of future payoffs for $\tau \geq \lambda$ is given by

$$\begin{aligned}
\overline{v}(\tau,\lambda) &= \frac{1}{2\delta}[w_C(\underline{\varepsilon})+w_C(\overline{\varepsilon})] \\
&\quad + \frac{(1-e^{-\lambda})[2+3\delta+e^{-\lambda}-\delta\frac{1-e^{-\lambda}}{1-e^{-\tau}}-(1+\delta)\Pi]}{4\delta[1+\delta(1-e^{-\tau})]}[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})]
\end{aligned}$$

9.2 Proof of Proposition 1

The first-order condition of $\overline{v}(0,\lambda)$ with respect to λ is given by

$$\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} = \left\{ -\frac{\delta(\Pi - 1 - e^{-\lambda})e^{-(2+\delta)\lambda}}{4\delta[1+\delta(1-e^{-\lambda})]^2} + \frac{(2+\delta)e^{-\lambda} - (1+\delta)(\Pi - 1)}{4\delta[1+\delta(1-e^{-\lambda})]} e^{-(1+\delta)\lambda} \right\} [w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]$$

Hence,

$$\begin{aligned} &\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} \geq 0 \\ &\Leftrightarrow -\delta(\Pi - 1 - e^{-\lambda})e^{-\lambda} \\ &\qquad + [(2+\delta)e^{-\lambda} - (1+\delta)(\Pi - 1)][1+\delta(1-e^{-\lambda})] \geq 0 \\ &\Leftrightarrow (\Pi - 1)\{\delta e^{-\lambda} + (1+\delta)[1+\delta(1-e^{-\lambda})]\} \\ &\qquad \leq \delta e^{-2\lambda} + (2+\delta)[1+\delta(1-e^{-\lambda})]e^{-\lambda} \\ &\Leftrightarrow \Pi \leq 1 + (1+\delta)\frac{2+\delta-\delta e^{-\lambda}}{(1+\delta)^2 e^{\lambda} - \delta^2}. \end{aligned} \tag{5}$$

Below, it is shown that the right-hand side of (5) unambiguously decreases in λ, taking on values in the range $[1, 2 + \frac{1}{1+2\delta}]$

$$\begin{aligned} \frac{\partial\left(\frac{2+\delta-\delta e^{-\lambda}}{(1+\delta)^2 e^{\lambda} - \delta^2}\right)}{\partial e^{-\lambda}} &= \frac{-\delta[(1+\delta)^2 e^{\lambda} - \delta^2] + e^{2\lambda}(1+\delta)^2[2+\delta-\delta e^{-\lambda}]}{[(1+\delta)^2 e^{\lambda} - \delta^2]^2} \\ &= \frac{\delta^3 + (1+\delta)^2 e^{\lambda}[(2+\delta)e^{\lambda} - 2\delta]}{[(1+\delta)^2 - \delta^2 e^{-\lambda}]^2} > 0. \end{aligned}$$

Since the right-hand side of (5) unambiguously decreases in λ, the following conclusion can be drawn: If $\Pi \leq 1$, then $\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} \geq 0$ for all $\lambda \geq 0$ and hence, $\widehat{\lambda} = \infty$. If $\Pi \geq 2 + \frac{1}{1+2\delta}$, then $\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} \leq 0$ for all $\lambda \geq 0$ and hence, $\widehat{\lambda} = 0$. If $\Pi \in (1, 2 + \frac{1}{1+2\delta})$, then there exists a unique $\widehat{\lambda} \in (0,\infty)$ such that $\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} \geq 0$ if and only if $\lambda \leq \widehat{\lambda}$, where $\widehat{\lambda}$ is determined by equalizing the left-hand side with the right-hand side of condition (5)

$$\Pi = 1 + (1+\delta)\frac{2+\delta-\delta e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2}$$

$$\Leftrightarrow (1+\delta)^2(\Pi-1)e^{\widehat{\lambda}} - \delta^2(\Pi-1) - (1+\delta)(2+\delta) + \delta(1+\delta)e^{-\widehat{\lambda}} = 0$$
$$\Leftrightarrow e^{-2\widehat{\lambda}} - \frac{\delta^2(\Pi-1)+(1+\delta)(2+\delta)}{\delta(1+\delta)}e^{-\widehat{\lambda}} + \frac{1+\delta}{\delta}(\Pi-1) = 0$$
$$\Leftrightarrow e^{-\widehat{\lambda}} = \frac{2+3\delta+\Pi\delta^2 \pm \sqrt{(2+3\delta+\Pi\delta^2)^2 - 4\delta(1+\delta)^3(\Pi-1)}}{2\delta(1+\delta)}$$

$$\Leftrightarrow \widehat{\lambda} = \ln[\frac{2\delta(1+\delta)}{2+3\delta+\Pi\delta^2 + \sqrt{(2+3\delta+\Pi\delta^2)^2 - 4\delta(1+\delta)^3(\Pi-1)}}]$$
$$\text{or } \widehat{\lambda} = \ln[\frac{2\delta(1+\delta)}{2+3\delta+\Pi\delta^2 - \sqrt{(2+3\delta+\Pi\delta^2)^2 - 4\delta(1+\delta)^3(\Pi-1)}}].$$

It is easily verified that the latter is the relevant solution (letting $\Pi \to 1$; it goes to infinity, while the former converges to $\ln(\frac{\delta}{2+\delta}) < 0$).

9.3 Proof of Lemma 1

Condition (5) from the previous proof can be rewritten as follows

$$\frac{\partial \overline{v}(0,\lambda)}{\partial \lambda} \geq 0 \Leftrightarrow \Pi \leq 1 + (1+\delta)\frac{2+\delta-\delta e^{-\lambda}}{(1+\delta)^2 e^{\lambda} - \delta^2}$$
$$\Leftrightarrow \Pi \leq 1 + e^{-\lambda}[1 + \frac{(1+\delta)e^{\lambda} - \delta}{(1+\delta)^2 e^{\lambda} - \delta^2}]. \qquad (5')$$

It is the case that

$$\frac{\partial[\frac{(1+\delta)e^{\lambda}-\delta}{(1+\delta)^2e^{\lambda}-\delta^2}]}{\partial\delta} = \frac{[(1+\delta)^2e^{\lambda}-\delta^2](e^{\lambda}-1) - 2[(1+\delta)e^{\lambda}-\delta]^2}{[(1+\delta)^2e^{\lambda}-\delta^2]^2}$$
$$= \frac{[(1+2\delta)e^{\lambda}+\delta^2(e^{\lambda}-1)](e^{\lambda}-1) - 2[e^{\lambda}+\delta(e^{\lambda}-1)]^2}{[(1+\delta)^2e^{\lambda}-\delta^2]^2}$$
$$= \frac{-\delta^2(e^{\lambda}-1)^2 + e^{\lambda}[(1-2\delta)(e^{\lambda}-1) - 2e^{\lambda}]}{[(1+\delta)^2e^{\lambda}-\delta^2]^2}$$
$$= \frac{-\delta^2(e^{\lambda}-1)^2 - e^{\lambda}[2+(1+2\delta)(e^{\lambda}-1)]}{[(1+\delta)^2e^{\lambda}-\delta^2]^2} < 0.$$

Hence, the right-hand side of (5') unambiguously decreases in δ. Since the right-hand side of (5') unambiguously decreases in λ, it immediately follows that λ will have to fall for the right-hand side not to change as δ rises. Thus, a higher δ will be associated with a lower $\widehat{\lambda}$. In particular, it follows that $\widehat{\lambda}$ may become zero as δ is increased.

9.4 Proof of Lemma 2

By introspection of expression (1), it is easily seen that $\overline{w}(\tau, \widehat{\lambda})$ is constant for $\tau \leq \widehat{\lambda}$. For $\tau > \widehat{\lambda}$, the first derivative is given by

$$\frac{\partial \overline{w}(\tau, \widehat{\lambda})}{\partial \tau} = -\frac{e^{-\tau}(1-e^{-\widehat{\lambda}})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]}{4(1-e^{-\tau})^2}[3 - \Pi - 2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}}].$$

It is easy to see that the sign of $\frac{\partial \overline{w}(\tau,\widehat{\lambda})}{\partial \tau}$ solely depends on the sign of $3-\Pi-2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}}$. Since $\Pi \in (1, 2+\frac{1}{1+2\delta})$, it is straightforward that $3 - \Pi - 2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}} < 0$ and hence, $\frac{\partial \overline{w}(\tau,\widehat{\lambda})}{\partial \tau} > 0$ for $\tau = \widehat{\lambda}$. As τ increases, $3 - \Pi - 2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}}$ increases monotonously, converging to

$$\begin{aligned}
\lim_{\tau\to\infty}(3 - \Pi - 2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}}) &= 1 - \Pi + 2e^{-\widehat{\lambda}} \\
&\overset{(5)}{=} 2e^{-\widehat{\lambda}} - (1+\delta)\frac{2+\delta-\delta e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2} \\
&= \frac{\delta(1+\delta)+\delta(1-\delta)e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2} > 0.
\end{aligned}$$

Hence, eventually $3 - \Pi - 2\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}}$ will become positive and $\frac{\partial \overline{w}(\tau,\widehat{\lambda})}{\partial \tau}$ will become negative. Thus, when $\widehat{\lambda} \in (0, \infty)$ is implemented, $\overline{w}$ initially increases and then decreases as τ increases beyond $\widehat{\lambda}$ under the optimal solution. The value to which it will converge is calculated as follows

$$\begin{aligned}
\lim_{\tau\to\infty} \overline{w}(\tau, \widehat{\lambda}) &= \frac{1}{2}w_C(\underline{\varepsilon}) + \frac{1}{2}w_C(\overline{\varepsilon}) + \frac{1-e^{-\widehat{\lambda}}}{4}(2+e^{-\widehat{\lambda}} - \Pi)[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\
&= \frac{1}{4}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\widehat{\lambda}}}{4}[\Pi - 1 - e^{-\widehat{\lambda}})][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]
\end{aligned}$$

$$\overset{(5)}{=} \frac{1}{4}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\widehat{\lambda}}}{4}[\frac{(1+\delta)(2+\delta-\delta e^{-\widehat{\lambda}})}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2} - e^{-\widehat{\lambda}}][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]$$

$$= \frac{1}{4}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\widehat{\lambda}}[1+\delta-\delta e^{-\widehat{\lambda}}]}{4[(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \quad (6)$$

$$> \frac{1}{4}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] = \overline{w}(0, \widehat{\lambda}).$$

9.5 Proof of Lemma 3

First, consider the case when $\tau \leq \widehat{\lambda}$. By introspection of expression (2), it is straightforward that $\overline{v}$ increases if and only if $\Pi \geq 1 + e^{-\widehat{\lambda}}$. For $\Pi \in (1, 2 + \frac{1}{1+2\delta})$, the optimal solution $\widehat{\lambda}$ is implicitly defined by equalizing both sides of condition (5), which in the proof of lemma 1 was rewritten as (5')

$$\Pi = 1 + e^{-\widehat{\lambda}}[1 + \frac{(1+\delta)e^{\widehat{\lambda}} - \delta}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2}] > 1 + e^{-\widehat{\lambda}}.$$

Hence, under the optimal solution, $\overline{v}$ increases unambiguously for $\tau \leq \widehat{\lambda}$. Next, consider the case when $\tau > \widehat{\lambda}$. For $\tau > \widehat{\lambda}$, the first derivative of $\overline{v}$ with regard to τ is given by

$$\frac{\partial \overline{v}(\tau, \widehat{\lambda})}{\partial \tau} = \frac{e^{-\tau}(1 - e^{-\widehat{\lambda}})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]}{4[1 + \delta(1 - e^{-\tau})]^2} \left\{ \begin{array}{l} [1 + \delta(1 - e^{-\tau})]\frac{1-e^{-\widehat{\lambda}}}{(1-e^{-\tau})^2} \\ -[2 + 3\delta + e^{-\widehat{\lambda}} - \delta\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}} - (1+\delta)\Pi] \end{array} \right\}$$
$$= \frac{e^{-\tau}(1 - e^{-\widehat{\lambda}})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]}{4[1 + \delta(1 - e^{-\tau})]^2} \left\{ \begin{array}{l} (1+\delta)(\Pi - 1) + 2\delta(\frac{1-e^{-\widehat{\lambda}}}{1-e^{-\tau}} - 1) \\ +\frac{1-e^{-\widehat{\lambda}}}{(1-e^{-\tau})^2} - (1 + e^{-\widehat{\lambda}}) \end{array} \right\}.$$

Hence,

$$\frac{\partial \overline{v}(\tau, \widehat{\lambda})}{\partial \tau} \geq 0 \Leftrightarrow (1+\delta)(\Pi - 1) + 2\delta(\frac{1 - e^{-\widehat{\lambda}}}{1 - e^{-\tau}} - 1) + \frac{1 - e^{-\widehat{\lambda}}}{(1 - e^{-\tau})^2} - (1 + e^{-\widehat{\lambda}}) \geq 0.$$

It is easy to see that the left-hand side decreases monotonously in τ. It equals

$(1+\delta)(\Pi-1)+\frac{e^{-2\widehat{\lambda}}}{1-e^{-\widehat{\lambda}}}>0$ for $\tau=\widehat{\lambda}$. Hence, $\overline{v}$ initially increases as τ increases beyond $\widehat{\lambda}$. The left-hand side decreases monotonously in τ, converging to

$$(1+\delta)(\Pi-1-2e^{-\widehat{\lambda}}) \overset{(5)}{=} \frac{1+\delta}{(1+\delta)^2e^{\lambda}-\delta^2}[-\delta(1+\delta)-\delta(1-\delta)e^{-\lambda}]<0.$$

Hence, $\overline{v}$ will initially increase and eventually decrease.

9.6 Proof of Proposition 4

For $\tau>\widehat{\lambda}$, the first derivative with regard to λ is given by

$$\frac{\partial\overline{v}(\tau,\lambda)}{\partial\lambda}=\frac{e^{-\lambda}[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})]}{4\delta[1+\delta(1-e^{-\tau})]}[(1+\delta)(1-\Pi)+2e^{-\lambda}+2\delta(1-\frac{1-e^{-\lambda}}{1-e^{-\tau}})].$$

Define $D\equiv(1+\delta)(1-\Pi)+2e^{-\lambda}+2\delta(1-\frac{1-e^{-\lambda}}{1-e^{-\tau}})$. It is straightforward that $\frac{\partial\overline{v}(\tau,\lambda)}{\partial\lambda}\geq 0$ if and only if $D\geq 0$. It is easy to see that D increases in τ and decreases in λ. Since $D>0$ for $\lambda=0$ and $D<0$ as $\lambda\to\infty$, it immediately follows that for any given τ, there exists a unique $\lambda\in(0,\infty)$ such that $D=0$ and hence $\overline{v}(\tau,\lambda)$ is maximized. Moreover, it follows that the solution to $\frac{\partial\overline{v}(\tau,\lambda)}{\partial\lambda}=0$ increases in τ. Hence, the ex post optimal solution will increase over time.

It remains to establish the relation of the ex post optimal adjustment phase length to the one being employed in the agreement. Setting $\lambda=\widehat{\lambda}$, the value for D when $\tau=\widehat{\lambda}$ is given by

$$\begin{aligned} D &= (1+\delta)(1-\Pi)+2e^{-\widehat{\lambda}} \\ &\overset{(5)}{=} 2e^{-\widehat{\lambda}}-(1+\delta)^2\frac{2+\delta-\delta e^{-\widehat{\lambda}}}{(1+\delta)^2e^{\widehat{\lambda}}-\delta^2} \\ &= \frac{-\delta(1+\delta)^2(1-e^{-\widehat{\lambda}})-2\delta^2e^{-\widehat{\lambda}}}{(1+\delta)^2e^{\widehat{\lambda}}-\delta^2}<0. \end{aligned}$$

Hence, the ex post optimal solution for the time limit will be smaller than the one agreed-upon as τ increases beyond $\widehat{\lambda}$. The value for D as $\tau\to\infty$ when $\lambda=\widehat{\lambda}$

is given by

$$\begin{aligned} D &= (1+\delta)(1-\Pi+2e^{-\widehat{\lambda}}) \\ &\overset{(5)}{=} (1+\delta)[2e^{-\widehat{\lambda}} - (1+\delta)\frac{2+\delta-\delta e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2}] \\ &= (1+\delta)\frac{\delta(1+\delta)+\delta(1-\delta)e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2} > 0. \end{aligned}$$

Hence, the ex post optimal solution for the time limit will eventually be larger than the one prescribed by the agreement.

9.7 Proof of Lemma 4

It must be investigated under what conditions $\lim_{\tau\to\infty}\overline{v}(\tau,\widehat{\lambda}) \geq \overline{v}(0,\widehat{\lambda})$. Using the optimality condition for $\widehat{\lambda}$, given by equalizing both sides of condition (5), the expression for $\overline{v}(0,\widehat{\lambda})$, given by (2), can be rearranged as follows

$$\begin{aligned} \overline{v}(0,\widehat{\lambda}) &= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] \\ &\quad + \frac{[\Pi-(1+e^{-\widehat{\lambda}})]e^{-(1+\delta)\widehat{\lambda}}}{4\delta[1+\delta(1-e^{-\widehat{\lambda}})]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\ &\overset{(5)}{=} \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] \\ &\quad + \frac{e^{-(1+\delta)\widehat{\lambda}}}{4\delta[1+\delta(1-e^{-\widehat{\lambda}})]}[\frac{(1+\delta)(2+\delta)-\delta(1+\delta)e^{-\widehat{\lambda}}}{(1+\delta)^2 e^{\lambda} - \delta^2} - e^{-\widehat{\lambda}}][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\ &= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-(1+\delta)\widehat{\lambda}}}{4\delta[(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]. \end{aligned}$$

Appropriately discounting the value to which $\overline{w}$ converges, given by (6), the following expression for the value to which $\overline{v}(\tau,\widehat{\lambda})$ converges is obtained

$$\begin{aligned} &\lim_{\tau\to\infty}\overline{v}(\tau,\widehat{\lambda}) \\ &= \frac{1}{4\delta}[2w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon}) + w_C(\overline{\varepsilon})] + \frac{e^{-\widehat{\lambda}}[1+\delta-\delta e^{-\widehat{\lambda}}]}{4\delta[(1+\delta)^2 e^{\widehat{\lambda}} - \delta^2]}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]. \end{aligned}$$

Since $1+\delta-\delta e^{-\widehat{\lambda}} > 1 > e^{-\delta\widehat{\lambda}}$ for $\widehat{\lambda} > 0$, it immediately follows that $\lim_{\tau\to\infty}\overline{v}(\tau,\widehat{\lambda}) <$

$\overline{v}(0, \widehat{\lambda})$ always holds. Hence, $\overline{v}(\tau, \widehat{\lambda}) < \overline{v}(0, \widehat{\lambda})$ for a sufficiently large τ.

9.8 Derivation of $v_N(\underline{\varepsilon})$ and $v_N(\overline{\varepsilon})$

The home country's value of the discounted flow of payoffs in the absence of any trade agreement v_N being in state $\underline{\varepsilon}$ and $\overline{\varepsilon}$ is given by

$$\begin{aligned}
v_N(\underline{\varepsilon}) &= \int_0^\infty [e^{-s} w_N(\underline{\varepsilon}) + (1 - e^{-s})\delta v_N(\overline{\varepsilon})] e^{-\delta s} ds \\
&= \frac{1}{1+\delta} w_N(\underline{\varepsilon}) + (\frac{1}{\delta} - \frac{1}{1+\delta})\delta v_N(\overline{\varepsilon}) \\
&= \frac{1}{1+\delta} [w_N(\underline{\varepsilon}) + v_N(\overline{\varepsilon})] \\
v_N(\overline{\varepsilon}) &= \int_0^\infty [e^{-s} w_N(\overline{\varepsilon}) + (1 - e^{-s})\delta v_N(\underline{\varepsilon})] e^{-\delta s} ds \\
&= \frac{1}{1+\delta} w_N(\overline{\varepsilon}) + (\frac{1}{\delta} - \frac{1}{1+\delta})\delta v_N(\underline{\varepsilon}) \\
&= \frac{1}{1+\delta} [w_N(\overline{\varepsilon}) + v_N(\underline{\varepsilon})].
\end{aligned}$$

Plugging the latter expression into the former yields

$$\begin{aligned}
v_N(\underline{\varepsilon}) &= \frac{1}{1+\delta} \{ w_N(\underline{\varepsilon}) + \frac{1}{1+\delta} [w_N(\overline{\varepsilon}) + v_N(\underline{\varepsilon})] \} \\
\Leftrightarrow [1 - \frac{1}{(1+\delta)^2}] v_N(\underline{\varepsilon}) &= \frac{1}{(1+\delta)^2} [(1+\delta) w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})] \\
\Leftrightarrow v_N(\underline{\varepsilon}) &= \frac{1}{\delta(2+\delta)} [(1+\delta) w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})].
\end{aligned}$$

Hence,

$$\begin{aligned}
v_N(\overline{\varepsilon}) &= \frac{1}{1+\delta} \{ w_N(\overline{\varepsilon}) + \frac{1}{\delta(2+\delta)} [(1+\delta) w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})] \} \\
&= \frac{1}{\delta(2+\delta)} [w_N(\underline{\varepsilon}) + (1+\delta) w_N(\overline{\varepsilon})].
\end{aligned}$$

9.9 Derivation of $v_C(\underline{\varepsilon})$ and $v_C(\overline{\varepsilon})$

Having been in state $\underline{\varepsilon}$ for $T \in [0, \lambda]$, the continuation value is given by

$$
\begin{aligned}
v_C(\underline{\varepsilon},T) &= \int_0^{\lambda-T}[e^{-s}w_N(\underline{\varepsilon})+(1-e^{-s})\delta v_C(\overline{\varepsilon})]e^{-\delta s}ds \\
&\quad + \int_{\lambda-T}^{\infty}[e^{-s}w_C(\underline{\varepsilon})+(1-e^{-s})\delta v_C(\overline{\varepsilon})]e^{-\delta s}ds \\
&= \int_0^{\infty}[e^{-s}w_C(\underline{\varepsilon})+(1-e^{-s})\delta v_C(\overline{\varepsilon})]e^{-\delta s}ds + \int_0^{\lambda-T}[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})]e^{-(1+\delta)s}ds \\
&= \frac{1}{1+\delta}w_C(\underline{\varepsilon})+(\frac{1}{\delta}-\frac{1}{1+\delta})\delta v_C(\overline{\varepsilon})+\frac{(1-e^{-(1+\delta)(\lambda-T)})}{1+\delta}[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] \\
&= \frac{1}{1+\delta}\{w_C(\underline{\varepsilon})+v_C(\overline{\varepsilon})+(1-e^{-(1+\delta)(\lambda-T)})[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})]\}.
\end{aligned}
$$

For $T > \lambda$, the continuation value is given by

$$
\begin{aligned}
v_C(\underline{\varepsilon},T) &= \int_0^{\infty}[e^{-s}w_C(\underline{\varepsilon})+(1-e^{-s})\delta v_C(\overline{\varepsilon})]e^{-\delta s}ds \\
&= \frac{1}{1+\delta}w_C(\underline{\varepsilon})+(\frac{1}{\delta}-\frac{1}{1+\delta})\delta v_C(\overline{\varepsilon}) \\
&= \frac{1}{1+\delta}[w_C(\underline{\varepsilon})+v_C(\overline{\varepsilon})].
\end{aligned}
$$

Being in state $\overline{\varepsilon}$, the continuation value is given by

$$
\begin{aligned}
v_C(\overline{\varepsilon}) &= \int_0^{\infty}[e^{-s}w_C(\overline{\varepsilon})+(1-e^{-s})\delta v_C(\underline{\varepsilon},0)]e^{-\delta s}ds \\
&= \frac{1}{1+\delta}w_C(\overline{\varepsilon})+(\frac{1}{\delta}-\frac{1}{1+\delta})\delta v_C(\underline{\varepsilon},0) \\
&= \frac{1}{1+\delta}[w_C(\overline{\varepsilon})+v_C(\underline{\varepsilon},0)].
\end{aligned}
$$

Setting $T = 0$ in the first expression and plugging it into the third expression yields

$$
\begin{aligned}
v_C(\overline{\varepsilon}) &= \frac{w_C(\overline{\varepsilon})}{1+\delta}+\frac{1}{(1+\delta)^2}\left\{\begin{array}{l} w_C(\underline{\varepsilon})+v_C(\overline{\varepsilon}) \\ +(1-e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] \end{array}\right\} \\
\Leftrightarrow\ [1-\frac{1}{(1+\delta)^2}]v_C(\overline{\varepsilon}) &= \frac{1}{(1+\delta)^2}\left\{\begin{array}{l} w_C(\underline{\varepsilon})+(1+\delta)w_C(\overline{\varepsilon}) \\ +(1-e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})] \end{array}\right\}
\end{aligned}
$$

$$\Leftrightarrow v_C(\overline{\varepsilon}) = \frac{1}{\delta(2+\delta)}\{w_C(\underline{\varepsilon}) + (1+\delta)w_C(\overline{\varepsilon}) + (1-e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\}.$$

Hence, having been in state $\underline{\varepsilon}$ for $T \in [0,\lambda]$, the continuation value is given by

$$\begin{aligned} v_C(\underline{\varepsilon},T) &= \frac{1}{1+\delta}\left\{ \begin{array}{l} w_C(\underline{\varepsilon}) + (1-e^{-(1+\delta)(\lambda-T)})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\ +\frac{w_C(\underline{\varepsilon})+(1+\delta)w_C(\overline{\varepsilon})+(1-e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon})-w_C(\underline{\varepsilon})]}{\delta(2+\delta)} \end{array} \right\} \\ &= \frac{1}{\delta(2+\delta)}\left\{ \begin{array}{l} (1+\delta)w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) \\ +[1+\delta - \frac{e^{-(1+\delta)\lambda}}{1+\delta} - \frac{\delta(2+\delta)e^{-(1+\delta)(\lambda-T)}}{1+\delta}][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \end{array} \right\}. \end{aligned}$$

For $T > \lambda$, the continuation value is given by

$$\begin{aligned} v_C(\underline{\varepsilon},T) &= \frac{w_C(\underline{\varepsilon})}{1+\delta} + \frac{1}{\delta(1+\delta)(2+\delta)}\left\{ \begin{array}{l} w_C(\underline{\varepsilon}) + (1+\delta)w_C(\overline{\varepsilon}) \\ +(1-e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \end{array} \right\} \\ &= \frac{1}{\delta(2+\delta)}\{(1+\delta)w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) + \frac{1-e^{-(1+\delta)\lambda}}{1+\delta}[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\}. \end{aligned}$$

9.10 Derivation of $v_C^*(\underline{\varepsilon})$ and $v_C^*(\underline{\varepsilon})$

If the home country has been in state $\underline{\varepsilon}$ for $T \in [0,\lambda]$, the foreign country's continuation value is given by

$$\begin{aligned} v_C^*(\underline{\varepsilon},T) &= \int_0^{\lambda-T} [e^{-s}w_N^* + (1-e^{-s})\delta v_C^*(\overline{\varepsilon})]e^{-\delta s}ds \\ &+ \int_{\lambda-T}^{\infty} [e^{-s}w_C^* + (1-e^{-s})\delta v_C^*(\overline{\varepsilon})]e^{-\delta s}ds \\ &= \int_0^{\infty} [e^{-s}w_C^* + (1-e^{-s})\delta v_C^*(\overline{\varepsilon})]e^{-\delta s}ds - \int_0^{\lambda-T} (w_C^* - w_N^*)e^{-(1+\delta)s}ds \\ &= \frac{1}{1+\delta}w_C^* + (\frac{1}{\delta} - \frac{1}{1+\delta})\delta v_C^*(\overline{\varepsilon}) - \frac{(1-e^{-(1+\delta)(\lambda-T)})}{1+\delta}(w_C^* - w_N^*) \\ &= \frac{1}{1+\delta}[w_C^* + v_C^*(\overline{\varepsilon}) - (1-e^{-(1+\delta)(\lambda-T)})(w_C^* - w_N^*)]. \end{aligned}$$

For $T > \lambda$, $v_C^*(\underline{\varepsilon},T)$ is given by

$$\begin{aligned} v_C^*(\underline{\varepsilon},T) &= \int_0^{\infty} [e^{-s}w_C^* + (1-e^{-s})\delta v_C^*(\overline{\varepsilon},0)]e^{-\delta s}ds \\ &= \frac{1}{1+\delta}w_C^* + (\frac{1}{\delta} - \frac{1}{1+\delta})\delta v_C^*(\overline{\varepsilon},0) \end{aligned}$$

$$= \frac{1}{1+\delta}[w_C^* + v_C^*(\overline{\varepsilon})].$$

The home country being in state $\overline{\varepsilon}$, the foreign country's continuation value is given by

$$\begin{aligned} v_C^*(\overline{\varepsilon}) &= \int_0^\infty [e^{-s} w_C^* + (1-e^{-s})\delta v_C^*(\underline{\varepsilon}, 0)] e^{-\delta s} ds \\ &= \frac{1}{1+\delta} w_C^* + (\frac{1}{\delta} - \frac{1}{1+\delta})\delta v_C^*(\underline{\varepsilon}, 0) \\ &= \frac{1}{1+\delta}[w_C^* + v_C^*(\underline{\varepsilon}, 0)]. \end{aligned}$$

Setting $T = 0$ in the first expression and plugging it into the third expression yields

$$\begin{aligned} & v_C^*(\overline{\varepsilon}) = \frac{1}{1+\delta} w_C^* + \frac{1}{(1+\delta)^2}[w_C^* + v_C^*(\overline{\varepsilon}) - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)] \\ & \Leftrightarrow [1 - \frac{1}{(1+\delta)^2}] v_C^*(\overline{\varepsilon}) = \frac{1}{(1+\delta)^2}[(2+\delta) w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)] \\ & \Leftrightarrow v_C^*(\overline{\varepsilon}) = \frac{1}{\delta(2+\delta)}[(2+\delta) w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)]. \end{aligned}$$

Hence, having been in state $\underline{\varepsilon}$ for $T \in [0, \lambda]$, the continuation value is given by

$$\begin{aligned} v_C^*(\underline{\varepsilon}, T) &= \frac{1}{1+\delta} \left\{ \begin{array}{l} w_C^* - (1 - e^{-(1+\delta)(\lambda - T)})(w_C^* - w_N^*) \\ + \frac{1}{\delta(2+\delta)}[(2+\delta) w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)] \end{array} \right\} \\ &= \frac{1}{\delta(2+\delta)}\{(2+\delta) w_C^* - (1 + \delta - \frac{e^{-(1+\delta)\lambda}}{1+\delta} - \frac{\delta(2+\delta) e^{-(1+\delta)(\lambda - T)}}{1+\delta})(w_C^* - w_N^*)\}. \end{aligned}$$

For $T > \lambda$, the continuation value is given by

$$\begin{aligned} v_C^*(\underline{\varepsilon}, T) &= \frac{1}{1+\delta}\{w_C^* + \frac{1}{\delta(2+\delta)}[(2+\delta) w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)]\} \\ &= \frac{1}{\delta(2+\delta)}[(2+\delta) w_C^* - \frac{1 - e^{-(1+\delta)\lambda}}{1+\delta}(w_C^* - w_N^*)]. \end{aligned}$$

9.11 Proof of Proposition 5

The ex ante expected continuation values under cooperation v_C^e and $\overline{v_C^*}$ equal the average of the continuation values under the two states for $T = 0$

$$\begin{aligned}
v_C^e &= \frac{1}{2}[v_C(\underline{\varepsilon}, 0) + v_C(\overline{\varepsilon})] \\
&= \frac{1}{2\delta}\{w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) + (1 - e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})]\} \\
v_C^{*e} &= \frac{1}{2}[v_C^*(\underline{\varepsilon}, 0) + v_C^*(\overline{\varepsilon})] \\
&= \frac{1}{2\delta}\{2w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*)\}.
\end{aligned}$$

Similarly, the home country's ex ante expected continuation value under the Nash equilibrium v_N^e equals the average of the continuation values under the two states for $T = 0$, while the foreign country's ex ante continuation value v_N^{*e} is given v_N^*

$$\begin{aligned}
v_N^e &= \frac{1}{2}[v_N(\underline{\varepsilon}, 0) + v_N(\overline{\varepsilon}, 0)] = \frac{1}{2\delta}[w_N(\underline{\varepsilon}) + w_N(\overline{\varepsilon})] \\
v_N^{*e} &= v_N^* = \frac{1}{\delta}w_N^*.
\end{aligned}$$

Hence,

$$\begin{aligned}
v_C^e - v_N^e &= \frac{1}{2\delta}\{w_C(\underline{\varepsilon}) + w_C(\overline{\varepsilon}) + (1 - e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] - w_N(\underline{\varepsilon}) - w_N(\overline{\varepsilon})\} \\
&= \frac{1}{2\delta}(\Pi - e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})] \\
v_C^{*e} - v_N^{*e} &= \frac{1}{2\delta}\{2w_C^* - (1 - e^{-(1+\delta)\lambda})(w_C^* - w_N^*) - 2w_N^*\} \\
&= \frac{1}{2\delta}(1 + e^{-(1+\delta)\lambda})(w_C^* - w_N^*).
\end{aligned}$$

The Nash bargaining product NBP is thus given by

$$\begin{aligned}
NBP &= (v_C^e - v_N^e)(v_C^{*e} - v_N^{*e}) \\
&= \frac{1}{4\delta^2}(\Pi - e^{-(1+\delta)\lambda})(1 + e^{-(1+\delta)\lambda})[w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})](w_C^* - w_N^*).
\end{aligned}$$

The first derivative with respect to λ is given by

$$\frac{\partial NBP}{\partial \lambda} = \frac{1}{4\delta^2} e^{-(1+\delta)\lambda} [2(1+\delta)e^{-(1+\delta)\lambda} - (1+\delta)(\Pi - 1)][w_N(\underline{\varepsilon}) - w_C(\underline{\varepsilon})](w_C^* - w_N^*).$$

Hence,

$$\begin{aligned} \frac{\partial NBP}{\partial \lambda} \geq 0 &\Leftrightarrow 2(1+\delta)e^{-(1+\delta)\lambda} - (1+\delta)(\Pi - 1) \geq 0 \\ &\Leftrightarrow \frac{1}{2}(\Pi - 1) \leq e^{-(1+\delta)\lambda}. \end{aligned}$$

The solution given by (4) immediately follows from this condition.

MONOGRAPH SERIES

1. **Michaely, Michael *The Theory of Commercial Policy: Trade and Protection*, 1973**
2. **Söderström, Hans Tson *Studies in the Microdynamics of Production and Productivity Change*, 1974**
3. **Hamilton, Carl B. *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost,* 1974**
4. **Nyberg, Lars and Staffan Viotti *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975**
5. **Myhrman, Johan *Monetary Policy in Open Economies*, 1975**
6. **Krauss, Melvyn *International Trade and Economic Welfare*, 1975**
7. **Wihlborg, Clas *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes*, 1976**
8. **Svensson, Lars E.O. *On Competitive Markets and Intertemporal Resources Allocation*, 1976**
9. **Yeats, Alexander J. *Trade Barriers Facing Developing Countries*, 1978**
10. **Calmfors, Lars *Prices, Wages and Employment in the Open Economy*, 1978**
11. **Kornai, János *Economics of Shortage*, Vols I and II, 1979**
12. **Flam, Harry *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981**
13. **Persson, Torsten *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach,* 1982**
14. **Erzan, Refik *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983**
15. **Horn af Rantzien, Henrik *Imperfect Competition in Models of Wage Formation and International Trade*, 1983**
16. **Nandakumar, Parameswar *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies,* 1985**
17. **Sellin, Peter *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990**
18. **Werner, Ingrid *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990**
19. **Svedberg, Peter *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991**
20. **Nordström, Håkan *Studies in Trade Policy and Economic Growth*, 1992**
21. **Hassler, John, Lundvik, Petter, Persson, Torsten and Söderlind, Paul *The Swedish Business Cycle: Stylized facts over 130 years*, 1992**
22. **Lundvik, Petter *Business Cycles and Growth*, 1992**
23. **Söderlind, Paul *Essays in Exchange Rates, Business Cycles and Growth*, 1993**
24. **Hassler, John A.A. *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994**
25. **Daltung, Sonja *Risk, Efficiency, and Regulation of Banks*, 1994**
26. **Lindberg, Hans *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994**
27. **Stennek, Johan *Essays on Information-Processing and Competition*, 1994**
28. **Jonsson, Gunnar *Institutions and Incentives in Monetary and Fiscal Policy*, 1995**
29. **Dahlquist, Magnus *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995**
30. **Svensson, Jakob *Political Economy and Macroeconomics: On Foreign Aid and Development,* 1996**
31. **Blix, Mårten *Rational Expectations and Regime Shifts in Macroeconometrics*, 1997**
32. **Lagerlöf, Nils-Petter *Intergenerational Transfers and Altruism*, 1997**
33. **Klein, Paul *Papers on the Macroeconomics of Fiscal Policy*, 1997**
34. **Jonsson, Magnus *Studies in Business Cycles,* 1997**
35. **Persson, Lars *Asset Ownership in Imperfectly Competitive Markets*, 1998**
36. **Persson, Joakim *Essays on Economic Growth*, 1998**
37. **Domeij, David *Essays on Optimal Taxation and Indeterminacy*, 1998**
38. **Flodén, Martin *Essays on Dynamic Macroeconomics*, 1999**
39. **Tangerås, Thomas *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000**
40. **Lidbom, Per Pettersson *Elections, Party Politics and Economic Policy*, 2000**

41. **Vestin, David *Essays on Monetary Policy*, 2001**
42. **Olofsgård, Anders *Essays on Interregional and International Political Economics*, 2001**
43. **Johansson, Åsa *Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity*, 2002**
44. **Groth, Charlotta *Topics on Monetary Policy*, 2002**
45. **Gancia, Gino A. *Essays on Growth, Trade and Inequality*, 2003**
46. **Harstad, Bård *Organizing Cooperation: Bargaining, Voting and Control*, 2003**
47. **Kohlscheen, Emanuel *Essays on Debts and Constitutions*, 2004**
48. **Olovsson, Conny *Essays on Dynamic Macroeconomics*, 2004**
49. **Stavlöt, Ulrika *Essays on Culture and Trade*, 2005**
50. **Herzing, Mathias *Essays on Uncertainty and Escape in Trade Agreements*, 2005**