ESSAYS IN INTERNATIONAL TRADE POLICY

by

Junhui Zeng

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

University of Wisconsin-Madison

2018

Date of final oral examination: 05/23/2018

The dissertation is approved by the following members of the Final Oral Committee:

Robert W. Staiger (Co-Chair), Professor of Economics, Dartmouth College

Kamran Bilir (Co-Chair), Assistant Professor of Economics

Lones Smith, Professor of Economics

Thomas Rutherford, Professor of Agricultural and Applied Economics

Acknowledgement

This dissertation would not have been possible without the advice and support of my advisors, other mentors, family and friends.

I would like to sincerely thank Robert W. Staiger, L. Kamran Bilir, and Lones Smith for their invaluable help and insightful comments. I am extremely grateful to Robert W. Staiger, for his continuous guidance and encouragement. I also thank John Kennan for the very helpful discussions.

Special thanks to my family for their untiring support. And I must also thank the friends and professors for the wonderful time I have during my graduate school.

Contents

Abstract

This dissertation consists of two self-contained chapters in international trade policy.

Chapter 1 theoretically and empirically explores bargaining delay in tariff negotiations. Imposing a bargaining structure in an equilibrium trade framework that captures a number of salient features of GATT multilateral tariff bargaining, I derive a number of predictions about delay in tariff negotiations. I then take a key prediction of my model to the newly compiled GATT (Torquay) tariff negotiation data, which includes detailed (product-level) information on the timing and contents of countries' (back-and-forth) proposals to each other. I find that my empirical exercise provides some support this prediction: delay in tariff negotiations is positively correlated with the magnitude of bargaining externalities, as measured (inversely) by the concentration of the exporting-country suppliers into the importing country's market.

The WTO is essentially a self-enforcing institution where cooperation partly hinges on members' ability to take actions or retaliate against deviations by others. However, in practice, the WTO does not allow for arbitrary retaliations once the punishment is triggered. In particular, the principle of reciprocity plays a role, which sets an upper bound for the retaliatory tariff rate. As a result, this restriction tends to reduce the pain from punishment, possibly making cooperation even harder. Chapter 2 examines why the WTO has such a restriction in place using a framework with imperfect private monitoring structure. Through a comparison of different retaliatory regimes, I argue that such restrictions indeed help member countries improve their welfare. It is further shown that under certain conditions countries choose to call upon the Panel ("court") only when a dispute actually occurs. Specifically, my analysis suggests that when

countries are sufficiently patient, a Call-upon regime for the Panel is the best among the three alternatives considered in the model.

Chapter 1. Bargaining Delay in Multilateral Trade Negotiations

1. Introduction

Delay in reaching agreements is not uncommon in international trade negotiations. For example, the most recent Doha Round of WTO negotiation is effectively in suspension, having been in process for about sixteen years. This is costly for countries in terms of welfare that could have been realized in a timely agreement. While there are many anecdotes regarding such delays in various media, a rigorous theoretical and empirical investigation of delays in international trade negotiations is, to my knowledge, missing. In this paper I take a first step toward filling this gap.

I focus on understanding delay in the context of multilateral rounds of tariff bargaining such as those undertaken in the GATT/WTO; and in particular I explore the possible contribution of bargaining externalities as a cause of delay in negotiations. GATT tariff negotiations provide an interesting laboratory within which to explore this possibility, for several reasons. First, GATT tariff negotiations are often carried out in the form of simultaneous bilateral bargains across many bilateral pairs of countries, with all tariff bargaining outcomes multilateralized to all member countries through the non-discrimination (most-favored-nation or MFN) principle of GATT: this means that GATT tariff bargains are typically settings of bilateral bargains with externalities. Second, the nature of the bargaining externality in this context is in principle quantifiable, as the externality travels through "world" prices and is hence subject to measurement. And third, the WTO has begun to declassify the detailed bargaining records of the earlier GATT rounds, providing an opportunity to explore the bargaining strategies implied by these real-world high-stakes bargaining records.

Bagwell, Staiger and Yurukoglu (2017) provide the exploration of the bargaining records from the GATT Torquay (1950-51) Round, and identify a set of stylized facts emerging from the bargaining records. As they observe, a number of these stylized facts point to a surprising lack of strategic behavior in GATT tariff bargains, a feature that has been noted by various commentators and GATT practitioners. Bagwell, Staiger and Yurukoglu argue that this feature can be interpreted as emerging from a tariff bargaining environment built around two pillars of the GATT architecture, namely, MFN (which requires that a county's import tariff on a given product must not vary by exporter source) and reciprocity, which in the GATT/WTO context is met when negotiated tariff changes lead to changes in a country's import volumes which are matched by changes in its export volumes. As these authors describe, reciprocity plays two roles in GATT: it is a norm for tariff negotiations aimed at achieving lower tariffs, and it defines a backup rule for tariff renegotiations that are aimed at raising tariffs from previously bound levels. As Bagwell and Staiger (forthcoming) formally demonstrate, when reciprocity in these two roles is joined with MFN and imposed as a set of strict constraints, the resulting bargaining forum acts like a posted-price mechanism under which countries have dominant strategies to make immediate tariff cut offers that reflect their true preferences. Bagwell, Staiger and Yurukoglu (2017) suggest that the GATT Torquay bargaining records can be interpreted through the lens of these findings.

In this paper, I follow Bagwell, Staiger and Yurukoglu (2017) in focusing on the bargaining records from the GATT Torquay Round. But I depart from their paper in two important ways. First, empirically, I explore the possibility that strategic dimensions of Torquay Round bargaining may be found in evidence on delayed offers, a dimension of the bargaining records that Bagwell, Staiger and Yurukoglu do not study. And second, to guide my empirical exploration of delay in tariff negotiations, from a theoretical perspective I depart from this prior work as well. Specifically, I maintain the GATT institutional focus on MFN and reciprocity in tariff negotiations, but I relax the strict insistence on reciprocity in *re*negotiations that Bagwell, Staiger and

Yurukoglu (2017) and Bagwell and Staiger (forthcoming) maintain.[1] As I demonstrate below, MFN plus reciprocity in tariff negotiations, but not in tariff renegotiations, leads to a bargaining environment where the terms-of-exchange (the "price") is fixed but the tariff cut depth (the volume of exchange) is negotiable. Further, I adopt an interpretation of the GATT principle supplier rule that implies a "rationing rule" that would not be allowed under the institutional assumptions maintained by Bagwell, Staiger and Yurukoglu and that provides further flexibility for negotiating the volumes of exchange.[2] In this institutional setting, with the terms of exchange fixed but the volume of exchange not fixed and hence negotiable, a conflict can result when the desires of the negotiating countries are not compatible, and strategic delay may arise.

To highlight the possible contribution of bargaining externalities as a cause of delay in negotiations, I consider within this institutional environment the offer of an MFN tariff cut from an importing country to two competing exporting countries who are each privately informed about the political pressures they face and who must collectively reciprocate the importing country's tariff-cut offer with tariff-cut offers of their own in order for the tariff bargain to succeed. The two exporting countries then make alternating offers about how to share the collective responsibility

[1]As Hoda (2001, pp. 83-87) details, reciprocity in (GATT Article XXVIII) renegotiation, as embodied in the right of a country to modify or withdraw a previous tariff concession subject to the reciprocal right of its principally affected negotiating partners to then withdraw reciprocal or "substantially equivalent" concessions, was present from the beginning of GATT in 1947, but (with certain exceptions described by Hoda) was available only 3 years (in the so-called "open season") after the initial negotiations had concluded. On the other hand, as Bagwell, Staiger and Yurukoglu (2017, footnote 27) note, the same reciprocity principle applies for temporary modifications of tariffs under GATT Article XIX, and there is no limitation on when such temporary modifications can be triggered. Reflecting these complexities, Bagwell, Staiger and Yurukoglu view their modeling of strict reciprocity in renegotiation as an approximation of a setting where such renegotiations could occur immediately upon conclusion of any negotiation. And viewed from this perspective, my relaxation of the reciprocity constraint in renegotiation can be viewed as an approximation to the alternative extreme where such renegotiations are not possible.

[2]As Bagwell and Staiger (forthcoming) and Bagwell, Staiger and Yurukoglu (2017) argue, when reciprocity in renegotiation is also imposed, a further "respect for voluntary exchange" constraint is added which, when combined with the class of rationing rules that these papers consider, leads to the dominant strategy result established in Bagwell and Staiger (forthcoming).

to reciprocate the importer country's offer. I show that in such a setting my model has a unique separating equilibrium in which bargaining delay may result, and that the length of delay will vary systematically with the strength of the bargaining externality across exporters, as measured by their exporter market concentration in the importing country's market.[3]

In particular, there are two (additive) components to the total amount of equilibrium bargaining delay in my model: the delay until an initial exporter offer is made, and the delay until the other exporter responds with an offer of its own.[4] Focusing first on the delay between the moment at which the initial exporter makes an offer and the time it takes for the second exporter to respond with an offer of its own and a final agreement to be reached, I show that this delay will be decreasing in the exporter market concentration into the importing country's market. Intuitively, this result arises in my model for two reasons. First, with private political pressure, an exporter may delay making an offer in order to credibly signal its type and hence the intensity of its aversion to a reciprocity-consistent offer that differs from its own ideal offer but would be required by reciprocity given the offer of the other exporter: in the central case that I consider, each exporting government may face either low or high political pressure from its export sector, and a government facing high pressure from its exporters will naturally value increased access and export volumes into the importing country market more highly than would a government facing low pressure, and may therefore wish to use delay to prevent mimicry by a low type and

[3]In particular, the country facing high political pressure from its export sector wants a relatively large increase in its export volume, compared with an otherwise symmetric country facing low political pressure. As a result, when the exporting countries compete for more trade volume - larger export tax cut/subsidy increase as in my model, they have the incentive to mimic the high pressure type. Then countries need to find ways to truthfully communicate their types. I find in the data there is observable delay between offers and counteroffers that vary significantly across deals, thus my model is one in which the length of delay can be used as a tool to signal type.

[4]As is standard in the literature (e.g. Admati and Perry 1987, and Harstad 2007), because each country truthfully signals its type in equilibrium, at most two offers are present - thus the maximum number of turns of delay is two which occurs when both countries are high type.

thereby signal its high type. And second, I assume that tariff bargaining follows the GATT/WTO principal supplier rule, under which an importing country chooses the principal (largest) exporter into its market as its primary bargaining partner for the tariff under consideration, and I interpret this rule as implying a natural ordering of offers in which the largest exporter makes the first offer and the smaller exporter then responds with its own offer. This means that the higher the exporter concentration, the smaller will be the exporter making the second offer in my model; and as I demonstrate, a smaller exporter need not delay as long in order to signal its type.

Turning next to the delay until the initial exporter offer is made, I find that this delay exhibits a non-monotonic relationship with exporter market concentration, first increasing and then decreasing. Intuitively, as exporter market concentration increases from an initially low level, the expected delay waiting for the second exporter country's response to the initial offer falls, and this increases the incentive of a low-type first exporter to mimic a high type and enjoy (now more quickly) the benefits of this mimicry, leading to the necessity of a longer delay in the initial offer of a high-type exporter in order to credibly signal its type. But for sufficiently high exporter market concentration the delay for the second exporter response goes to zero, and hence further increases in exporter market concentration no longer have any impacts on first-offer delay through this channel; and eventually the exporter market concentration becomes high enough that the second exporter has a trivial impact on the collective reciprocal response of the two exporters to the importer, and the incentive of a low-type first exporter to mimic a high type then begins to fall, leading to the necessity of shorter delay in the initial offer of a high-type exporter in order to credibly signal its type.

Putting these two components of equilibrium delay together, my model delivers the surprising result that overall delay is non-monotonic with respect to exporter market concentration, and hence with respect to the magnitude of the bargaining externality in my model. Nevertheless, as this overall non-monotonicity reflects the sum of two delay components, one of which *does* exhibit a monotonic relationship with exporter market concentration, my model delivers a simple prediction that can be taken to the GATT bargaining data.[5]

My model also provides a novel rationale for the GATT principal supplier rule, which generally advises countries to focus their negotiation on a given product with principal (largest) supplier(s) of that product. Since the discussion of this feature requires some fundamental asymmetry in the model, from an *ex ante* perspective, a framework with ex ante symmetry between players (e.g., Harstad, 2007) does not offer an appropriate way to evaluate this kind of rule. Based on a setup with a more developed economic structure, for example, one with an explicit demand/supply system, my setting not only provides a way to evaluate the asymmetry, but also delivers a natural interpretation of the benefits of the principal supplier rule. In particular, under certain plausible conditions, I am able to show that this rule will shorten the amount of overall bargaining delay in the model. Thus, from this perspective, the theory lends some support to the principal supplier rule also embodied in GATT articles.

As described above, my model delivers a simple prediction that can be taken to the GATT bargaining data: a particular component of bargaining delay should be decreasing in exporter market concentration. More specifically, my model predicts that the delay from the first exporter offer to the final agreement is decreasing in the Herfindahl-Hirschman index (HHI) of exporter concentration into the import market, where the HHI is calculated over the bargaining exporters but excluding the export shares of non-bargainers (as I discuss below, this last feature follows from a property

[5]In stating this, I am focusing on the central case where the home country is in modestly short position, although my empirical work accounts for both the home long and the home short case.

of tariff negotiations satisfying MFN and reciprocity that is emphasized by Bagwell and Staiger, 2005, forthcoming, and Bagwell, Staiger and Yurukoglu, 2017).

I focus on the tariff bargaining records from the 1950-51 GATT Torquay Round as compiled and described in Bagwell, Staiger and Yurukoglu (2017). My model has only two (bargaining) exporters, but as I later confirm importing countries rarely negotiate with more than two exporting countries on any given tariff in the data. To take my model prediction to this data, I focus primarily on US tariff offers, where product level bilateral import data for the period is also available as compiled and described by Bagwell, Staiger and Yurukoglu (2015), and hence where the HHI can be calculated. I identify the home country in my model with the US, and I identify the foreign exporters bargaining with the home country in my model over the tariff on a given product as the foreign exporters that received a tariff-cut offer from the US on that product (as revealed by the bargaining records). This allows me to calculate the HHI of exporter concentration that the US faced when bargaining over each tariff cut that it offered in the Torquay Round, which according to my model is an inverse measure of the bargaining externality associated with the negotiations on each particular US tariff cut. The bargaining records also record the dates of each offer and counter-offer, allowing me to calculate the delay from the first offer received by the US from a foreign exporter to which it had made a tariff cut offer on a particular product to the time of the final agreement by the US to cut that tariff.

Making use of these measures, my empirical findings provide some evidence of a positive relationship between bargaining externalities and delay for the US tariff offers at Torquay, offering some support for a key prediction of my model. Furthermore, while my main empirical results focus on the US tariff offers where the available trade data allow me to construct an externality measure that conforms closely with my theory, I also confirm the findings with the entire set of Torquay Round tariff offers when I use a proxy for the bargaining externality the construction of which does not require trade data.

The remainder of the paper is organized as follows. I review the related literature in Section 2. An overview of the theory is provided in Section 3 based on a general equilibrium setting, to illustrate the motives of countries to bargain at all. In Section 4, a formal model is introduced, containing the key structure of the theory. I conduct the equilibrium analysis in Section 5. Using my modeling results as a guide, Section 6 explores the relationship between bargaining delay and exporter market concentration in the context of the GATT Torquay Round bargaining records. Section 7 concludes. Unless otherwise noted, all proofs are provided in the Appendix.

2. Related Literature

In this section, I provide a brief discussion of the related literature. Broadly speaking, bargaining delay has attracted enormous interest among bargaining theorists, at least since Rubinstein's seminal work on "splitting the pie", which however predicts an immediate agreement between two players. Starting from this basic setting, the literature on bargaining delay can be divided into two categories: bargaining delay due to informational frictions (incomplete information or/and imperfect information) or bargaining delay due to various bargaining externalities. In the strand of literature on bargaining with private information, players have private evaluations of the underlying good in the bargaining game. In this literature, (costly) delay is a possible way to truthfully reveal their evaluation:[6] in general, the longer a player is willing to delay an agreement, the stronger its bargaining strength will be (e.g. Admati and Perry 1987).

[6]The cost of delay could come from discounting or the increased possibility of bargaining failure. In the GATT setting, discounting could be interpreted as being related to the probability of an exogenous end to the round in any period, so the risk of delay is that countries might not reach agreement before the exogenous end.

Compared with the vast literature on bargaining with incomplete information, the literature on bargaining with externalities is relatively sparse. This strand of the literature usually assumes that the underlying object has the property of being a "public good." Jehiel and Moldovanu (1995a) examine a finite-horizon framework, where a seller bargains with many potential buyers, and the (abstract and exogenous) non-pecuniary externality leads to delay as every buyer has an incentive to wait for others to bear the cost of the good. Yet the delay result crucially depends on the matching friction where the seller can only be randomly matched to a buyer in each given round. An extension to infinite horizon is considered in Jehiel and Moldovanu (1995b), where "cyclical delay" emerges, although a complete characterization of equilibria is missing, given the complexity. There is also a line of literature that focuses on coalitional bargaining with externalities that can result in delay. Gomes (2005) analyzes an abstract coalitional game where multiple players are deciding the membership of the coalition. In this setting, delay is possible because of players' randomization. Thus, while delay can be generated in these settings, the prediction is usually not sharp, as typically there are multiple equilibria.

A paper that takes an approach closely related to mine is Harstad (2007). In Harstad's setting, both informational frictions and externalities are incorporated. In particular, in Harstad's model the presence of private information is key to the existence of delay, while the magnitude of externality determines the length of delay. However, delay is simply an intermediate instrument in Harstad's model that serves to evaluate the effects of equalization and side payments on countries' ex ante welfare.[7] As a result, Harstad does not focus on analyzing delay as I do.

[7]Indeed, as pointed out by Harstad (2007), the evaluation could be otherwise done without the use of delay.

There are important features of both private information and externalities in the GATT/WTO trade bargaining setting, but the existing bargaining models don't reflect well the way these features arise in the trade bargaining setting - and usually they are too abstract to be taken to the data. In the GATT/WTO tariff bargaining setting, a given importer with a (nondiscriminatory) tariff cut to offer often negotiates with several foreign exporters who in a successful bargain must then each make tariff offers back to the importer that together reciprocate the trade liberalization offered by the importer. While the analogue to the underlying object that players bargain over in the typical bargaining model (e.g. the seller-buyer bargaining games, Rubinstein's pie-splitting game) is the amount of trade liberalization that is needed to reciprocate the importer's trade liberalization offer, a key difference between the typical bargaining model and the tariff bargaining setting just described is clear: in the typical bargaining model, the bliss point for each player is either the whole object or none of the object, but in the tariff bargaining setting the bliss point for each foreign exporter is the ideal tariff cut the exporter would like to offer if it knew that its offer together with the offers of other exporters would reciprocate the tariff cut offer made by the importing country. Moreover, in the tariff bargaining setting this bliss point will typically be interior (i.e., it will not typically correspond to either a fully reciprocating tariff cut – "the whole pie" – or no tariff cut – "none of the pie") and will depend on the political preferences of the exporting country, which may be private information to the exporting government and therefore unknown to the other exporting governments involved in the bargain. This feature complicates the translation of the typical bargaining setting into the GATT/WTO tariff bargaining setting. In particular, countries may compete for making concessions whether they are collectively competing for more liberalization or less liberalization.

Another way in which the tariff bargaining setting departs from the standard bargaining models in the literature is that, while externalities are introduced with an *exogenous* reduced-form assumption in the typical bargaining models and are therefore interpreted in an abstract fashion, in the GATT/WTO tariff bargaining setting that I analyze here the bargaining externality is *endogenous*, being an intrinsic feature of the economic environment and institutional rules. This departure from standard bargaining settings has two important implications.

A first implication is that a more detailed modeling of the externality is required in the tariff bargaining setting. Specifically, the MFN restriction of nondiscrimination requires GATT/WTO members to apply uniform tariffs to all other members on a given product, and this ensures that all exporters, whether or not they engage in tariff bargaining, will face the same conditions of access into a given import market as summarized by a common-across-exporters world price. Hence, when there are multiple countries exporting to a given import market, each exporting country has an incentive to free ride on the bargaining efforts of the other exporting countries, and the observable export bilateral volumes serve as a natural candidate for measuring the strength of the externalities in the trade bargaining: at the extreme, if only one country exports to a given importing country, then no free-riding externality problem should exist in bargaining over that import tariff. The reciprocity restriction, when coupled with MFN, further shapes this externality problem and reduces it to the aforementioned bliss point issue: each exporter into a given import market would like the competing exporters into that market to be the residual claimants on any tariff liberalization needed to reciprocate the tariff cut on offer; that is, each exporter would like the other competing exporters to agree to tariff cuts which move them away from their bliss points in order to reciprocate the importer tariff cut on offer, so that it can achieve its own bliss point. This implies that each country could be either competing with other exporting countries to cut its tariff more than is jointly compatible with the reciprocation of the importer tariff cut on offer, or competing to cut its tariff less than is jointly compatible with reciprocation of the importer tariff cut on offer.

A second implication of this departure from standard bargaining settings is that, rather than taking an exogenous and abstract form, the externality in the GATT/WTO tariff bargaining setting is readily interpretable in terms of model primitives. This has the advantage that the externality in the tariff bargaining setting is measurable, and its relation to bargaining delay can in principle be examined in the data.

There are also a number of additional interesting features that distinguish the trade bargaining from those in classical theories. (i) While this traditional line of literature generally focuses on linear structures (e.g. buyer/seller splitting the residual from trade via prices), the trade bargaining usually involves the consideration of each country's bliss point (thus a non-linear structure), which may or may not be compatible with each other.[8] That said, as I noted above, countries may compete to make concessions, but they also compete to avoid making concessions. (ii) The source of incomplete information can take many forms, and come from many fundamental parameters that are not observable by others. (iii) In an ideal framework, trade bargaining involves many players and many goods.

3. Theory: Overview

To illustrate the bargaining problem, I start my analysis within general equilibrium frameworks, by considering two countries first to demonstrate the basic features of tariff bargaining under reciprocity in tariff negotiations – but without the possibility of renegotiation under reciprocity as Bagwell and Staiger (forthcoming) and Bagwell, Staiger and Yurukoglu (2017) consider. And I then consider three countries to exhibit the bargaining under both reciprocity and MFN.

3.1. **Bargaining under Reciprocity.** To illustrate at a general level the basic features of tariff bargaining under reciprocity that are at the heart of the theoretical

[8]By "compatible with each other", I mean that the bliss points would allow both exporters to achieve their respective bliss points while exactly reciprocating together the tariff cut on the offer from the importer.

and empirical analysis to follow, I start by considering just two countries, and in particular begin with a standard two country - two good general equilibrium model. I refer to the two countries as the home country (whose variables are without a superscript $\star$), and the foreign country (whose variables are with a superscript $\star$), respectively. In this setting, the home country is a natural importer of good x, with the foreign country being a natural importer of y, and given the (import) tariffs $(\tau^\star, \tau)$ of the countries, the equilibrium world relative price can be written as $p^w(\tau^\star, \tau)$. Furthermore, the local relative prices in the home and the foreign markets are $p \equiv p^w(1+\tau)$ and $p^\star \equiv p^w(1+\tau^\star)$, respectively. Given countries' (general) welfare functions $W(\tau, \tau^\star)$ and $W^\star(\tau^\star, \tau)$,[9] the Nash equilibrium $(\tau^\star_N, \tau_N)$ implies a world price $p^w_N(\tau^\star_N, \tau_N)$. As established in Bagwell and Staiger (1999), the welfare functions can be alternatively represented as $W(\tau, \tau^\star) \equiv \widehat{W}(p(\tau, p^w), p^w(\tau^\star, \tau))$ and $W^\star(\tau^\star, \tau) \equiv \widehat{W}^\star(p(\tau^\star, p^w), p^w(\tau^\star, \tau))$, and reciprocity implies that the world price should be unchanged, whatever the agreement would be.[10] In particular, under this constraint, if the existing tariffs are Nash, then any agreement, in the $(\tau^\star, \tau)$ domain, must lie on the iso-world-price locus $\{(\tau^\star, \tau) | p^w(\tau^\star, \tau) = p^w_N(\tau^\star_N, \tau_N)\}$. Then, the question becomes which portion along this locus constitutes the bargaining frontier. And, as I will argue later, this characterization of the bargaining frontier is the key that distinguishes my work from Bagwell, Staiger and Yurukoglu (2017).

Consider the ideal tariff (and hence the ideal local price) that each country would want, while the $p^w(\tau^\star, \tau)$ was maintained at $\overline{p^w}$ by the needed change from its trading partner. Formally, for the home country, its preferred tariff pair is the $(\tau, \tau^\star)$ that

[9]As pointed out in Bagwell and Staiger (1999,2002), a general payoff function can be used to nest many popular models of trade policies, whether countries are politically motivated or not. For example, $W^j(p^j, p^w)$ may be the sum of consumer surplus and producer surplus, where a politically motivated government might put higher weights on the producer surplus. Thus, any tariff change made by country j will affect its payoff through its effects on the local price p^j and the world price p^w, where the latter is often regarded as the 'terms-of-trade externality' in the literature.

[10]As Bagwell and Staiger (1999) show, this follows from trade balance as long as reciprocity is defined as changes in tariffs that lead to changes in the volume of each country's imports that are of equal value to the changes in the volume of its exports.

solves $W_p \equiv \widehat{W}_p(p(\tau, \overline{p^w}), \overline{p^w}) = 0$ and $p^w(\tau^\star, \tau) = \overline{p^w}$; for the foreign country, it is the $(\tau, \tau^\star)$ that solves $W^\star_{p^\star} \equiv \widehat{W}^\star_{p^\star}(p^\star(\tau^\star, \overline{p^w}), \overline{p^w}) = 0$ and and $p^w(\tau^\star, \tau) = \overline{p^w}$.[11] Thus, the only relevant portion for bargaining is the interval, contained by $\{(\tau^\star, \tau) | p^w(\tau^\star, \tau) = \overline{p^w}\}$, with two endpoints being the ideal tariff pairs of the two countries. That is, the bargaining frontier is the part of the iso-world-price locus above the lowest of these pairs and below the highest of these pairs.

Figure 3.1 can be used to illustrate this setup. In this figure, all the indifference curves and iso-world-price locus are represented in the $(\tau^\star, \tau)$ domain. Depending on the setting, the initial Nash equilibrium could be denoted by $N(A)$, $N(B)$ or $N(C)$, and reciprocity in negotiation means countries' bargaining outcome has to lie on the iso-world-price $p^w_N(\cdot)$ lines going through them, where the terms of exchange is fixed by the reciprocity restriction. In particular, this can lead to three cases: (1) If the existing Nash equilibrium is $N(A)$, then countries have to bargain along the $p^w_N(A)$ line, where the bargaining frontier is characterized by the segment $A'A''$. In this case, the home country is *long*, in the sense that it wants more liberalization (A'') than the foreign country wants (A'). (2) If the existing Nash equilibrium is $N(B)$, then countries have to bargain along the $p^w_N(B)$ line, where the bargaining frontier is characterized by the segment $B'B''$. In this case, the home country is *short*, in the sense that it wants less liberalization (B') than the foreign country wants (B''). (3) If the existing Nash equilibrium is $N(C)$, then countries have to bargain along the $p^w_N(C)$ line which is equal to the world price under political optimum - the bargaining frontier simply reduces to a single point PO, which is also the political optimum. In this case, the wants of the two countries are exactly matched. I just denote this as the case where the home country is *equal*.

[11] These tariffs (the own-tariff component of the ideal tariff pair for each country) are what Bagwell and Staiger (1999) define as the *politically optimal* reaction curve tariffs.

As a comparison, Bagwell, Staiger and Yurukoglu (2017) impose the additional restriction of reciprocity in *re*negotiation under which no country could be forced to import more than they want. As a result of this “respect for voluntary exchange” constraint, countries’ bargaining frontiers in their model are necessarily subsets of the frontiers in mine. In particular, reciprocity in renegotiation prescribes that any agreement must also lie on $W^{\star}_{p^{\star}} = 0$ or $W_p = 0$, whichever yields smaller trade volume. Thus, together with reciprocity in negotiation, this additional restriction pins down a unique negotiation outcome, that is its intersection with the iso-world-price locus: the bargaining frontier in Bagwell, Staiger and Yurukoglu is therefore a single point. In particular, in the three cases described previously and in terms of Figure 3.1, the bargaining frontiers in their setting are A', B' and PO respectively. Thus, countries effectively do *not* need to bargain at all, as Bagwell, Staiger and Yurukoglu observe.

While the bargaining frontier is the same single point in the knife-edge case where the existing world price coincides with the world price under political optimum, however, without reciprocity in renegotiation, reciprocity in negotiation alone requires countries to determine a point on the line segment $A'A''$, $B'B''$ in more general cases. Thus, bargaining over the trade volume (though not the terms of trade) arises, which is exactly my starting point.

3.2. **Bargaining under Reciprocity and MFN.** The two-country setting above is useful for illustrating the basic features of reciprocity in negotiation and in renegotiation, and how each impacts the bargaining frontier. But to consider the possibility of bargaining externalities and the impact of the MFN rule together with reciprocity in shaping these externalities, I must move to a many-country framework. A simplest but natural framework to analyze MFN and Reciprocity involves three countries. Again I work within a simple general equilibrium setting to illustrate the key idea behind my analysis. Here I follow the setup described in Bagwell and Staiger (1999,

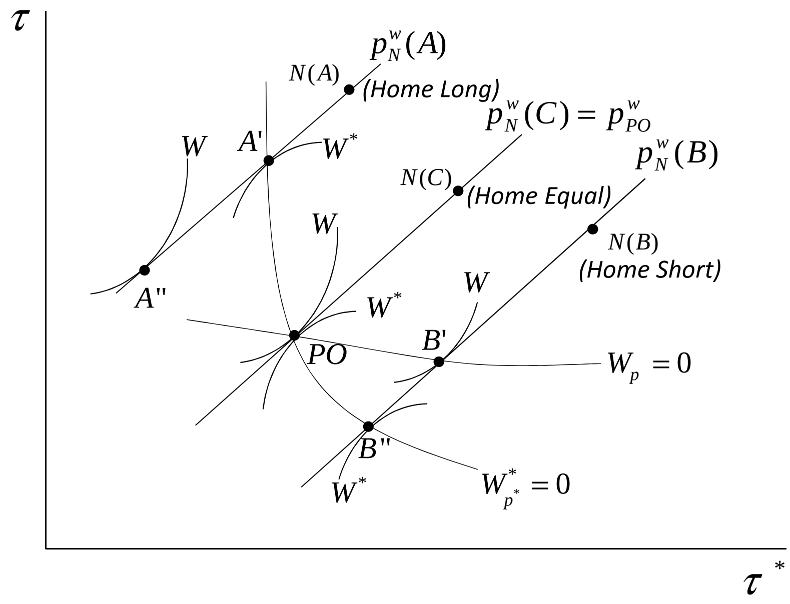

FIGURE 3.1.

2002). In particular, I assume that there are two goods, x and y, in this setting. Also, among the three countries, there is one home country (whose variables are without a superscript) that is a natural importer of good x from two foreign countries, denoted by $\star j$, with $j \in \{1, 2\}$, which are natural importers of good y from the home country. I denote the local relative price in the home country as $p \equiv \frac{p_x}{p_y}$, and the local relative price in the foreign country as $p^{\star j} \equiv \frac{p_x^{\star j}}{p_y^{\star j}}$ for $j \in \{1, 2\}$. Assuming all countries use non-prohibitive ad valorem import taxes, with the home country's (foreign countries') MFN import tariff denoted by τ ($\tau^{\star j}$), then the world relative price can be written as $p^w \equiv \frac{p_x^\star}{p_y}$, where $p_x^\star \equiv p_x^{\star 1} = p_x^{\star 2}$ and local prices are such that $p(\tau, p^w) = p^w \cdot (1 + \tau)$ and $p^{\star j}(\tau^{\star j}, p^w) = p^w/(1 + \tau^{\star j})$.[12]

I denote the home country's domestic demand of good i as $D_i(p, TR)$ and the foreign country's demand as $D_i^{\star j}(p^{\star j}, TR^{\star j})$, $i \in \{x, y\}$, where TR and $TR^{\star j}$ respectively represents the tariff revenue of the home country and the foreign countries,

[12]Under discriminatory tariffs, there will be a world relative price for each foreign country. However, the MFN restriction, together with law of one price, requires the world relative prices are the same across the home country's two trading partners.

which I assume is a lump-sum transfer to its own domestic consumers. Also, letting the home country's domestic supply be $Q_i(p)$ and the foreign country's supply be $Q_i^{\star j}(p^{\star j})$, then the net import of good x by the home country, and that of y by $\star j$ is respectively $M_x(p, TR) \equiv D_x(p, TR) - Q_x(p)$ and $M_y^{\star j}(p^{\star j}, TR^{\star j}) \equiv D_y^{\star j}(p^{\star j}, TR^{\star j}) - Q_y^{\star j}(p^{\star j})$. Similarly, the net export of good y by the home country, and that of x by $\star j$ is respectively $E_y(p, TR) \equiv Q_y(p) - D_y(p, TR)$ and $E_x^{\star j}(p^{\star j}, TR^{\star j}) \equiv Q_x^{\star j}(p^{\star j}) - D_x^{\star j}(p^{\star j}, TR^{\star j})$. Thus TR and $TR^{\star j}$, measured in units of each country's export good, is implicitly given by $TR = \tau \cdot p^w M_x(p, TR) = (p - p^w) M_x(p, TR)$ and $TR^{\star j} = \tau^{\star j} \cdot \frac{1}{p^w} M_y^{\star j}(p^{\star j}, TR^{\star j}) = (\frac{1}{p^{\star j}} - \frac{1}{p^w}) \cdot M_y^{\star j}(p^{\star j}, TR^{\star j})$, which yields $TR = TR(p(\tau, p^w), p^w)$ and $TR^{\star j} = TR^{\star j}(p^{\star j}(\tau^{\star j}, p^w), p^w)$. Thus, given the world price p^w and the set of local prices: $p(\tau, p^w), p^{\star 1}(\tau^{\star 1}, p^w), p^{\star 2}(\tau^{\star 2}, p^w)$, consumption, production and import/export in each country will be entirely pinned down. And countries' tariff choice $(\tau, \tau^{\star 1}, \tau^{\star 2})$ determines such prices in equilibrium.[13]

The equilibrium world relative price, denoted by p_{eq}^w, is given by the market clearing condition in the x market:[14]

$$M_x(p(\tau, p_{eq}^w), TR(p(\tau, p_{eq}^w), p_{eq}^w)) =$$

$$E_x^{\star 1}(p^{\star 1}(\tau^{\star 1}, p_{eq}^w), TR^{\star 1}(p^{\star 1}(\tau^{\star 1}, p_{eq}^w), p_{eq}^w)) + E_x^{\star 2}(p^{\star 2}(\tau^{\star 2}, p_{eq}^w), TR^{\star 2}(p^{\star 2}(\tau^{\star 2}, p_{eq}^w), p_{eq}^w))$$

Thus, p_{eq}^w can be written as $p_{eq}^w(\tau, \tau^{\star 1}, \tau^{\star 2})$. Furthermore, I assume that the Lerner paradox is ruled out, such that $\frac{\partial p_{eq}^w(\tau, \tau^{\star 1}, \tau^{\star 2})}{\partial \tau^h} < 0 < \frac{\partial p_{eq}^w(\tau, \tau^{\star 1}, \tau^{\star 2})}{\partial \tau^{\star j}}$. That is, the world price is now decreasing in the home tariff and increasing in each of the foreign tariffs. And, because of this condition, given an initial $p_{eq}^w(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$, to reciprocate any tariff cut $(\tau_0 - \tau > 0)$ offered by the home country, some combination of tariff cuts

[13]Presumably, countries may impose both export and import policies. However, I have to restrict the policy choice in a general equilibrium setting, because Lerner Symmetry would imply that there will be indeterminacy in the equilibrium if both export taxes and import taxes are allowed, that is, there are many combinations of them that yield the same outcome.

[14]Market clearing for good y is redundant, when x good market is cleared and budget constraints are satisfied - ensured by balanced trade.

$(\tau_0^{\star 1} - \tau^{\star 1} > 0, \tau_0^{\star 2} - \tau^{\star 2} > 0)$ have to be granted by the two foreign countries such that $p^w_{eq}(\tau, \tau^{\star 1}, \tau^{\star 2}) = p^w_{eq}(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$.

Again, this leads to the possibilities of *home long*, *home short* and *home equal* (*knife-edge*) as I discussed in the previous two-country setting. To illustrate this in a three country setting, I assume countries' payoff can be represented by $W^j(p^j, p^w)$ for $j \in \{\text{none}, \star 1, \star 2\}$ where "none" means no superscript. Because of my focus on the three country model from now on, I will explicitly express the determination of countries' reaction curves. First, the standard terms-of-trade effect can be written as:

$$\frac{\mathrm{d}W^j(p^j, p^w)}{\mathrm{d}\tau^j} = \frac{\partial W^j(p^j, p^w)}{\partial p^j}\frac{\mathrm{d}p^j(\tau^j, p^w_{eq})}{\mathrm{d}\tau^j} + \frac{\partial W^j(p^j, p^w)}{\partial p^w}\frac{\partial p^w_{eq}}{\partial \tau^j}$$

Thus, other things equal, if the government values its terms-of-trade improvement, that is $\frac{\partial W(p,p^w)}{\partial p^w} < 0$ and $\frac{\partial W^j(p^j,p^w)}{\partial p^w} > 0$ for $j \in \{\star 1, \star 2\}$, and if it wants to use its trade policy to achieve a preferred local price, then it has the incentive to set its import tariff and shift the cost associated with this tariff to its trading partners. Moreover, each country's reaction curve can be characterized by $\frac{\mathrm{d}W^j(p^j,p^w)}{\mathrm{d}\tau^j} = 0$. In a Nash Equilibrium, each country stays on its reaction curve, leading to an inefficient outcome in general.

Under the rule of reciprocity (i.e. $\mathrm{d}p^w_{eq} = 0$), $\frac{\mathrm{d}W^j(p^j,p^w)}{\mathrm{d}\tau^j} = 0$ is equivalent to $\frac{\partial W^j(p^j,p^w)}{\partial p^j} = 0$. Thus, as in the two-country setting, each country's preferred tariff is again characterized by $W^j_{p^j} = 0$, under the condition that the world relative price will be maintained. And, for notational convenience, I will denote any variable that maximizes a country's payoff under fixed world price with a subscript "*po*" throughout the paper. In particular, the home country prefers τ_{po}, pinned down by $W_p = 0$, given that it will be reciprocated by some $(\tau_r^{\star 1}, \tau_r^{\star 2})$ such that $p^w_{eq}(\tau_{po}, \tau_r^{\star 1}, \tau_r^{\star 2}) = p^w_{eq}(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$. Foreign country $\star 1$ prefers $\tau^{\star 1}_{po}$, pinned down by $W^{\star 1}_{p^{\star 1}} = 0$, given that it will be reciprocated by some $(\tau_r, \tau_r^{\star 2})$ such that $p^w_{eq}(\tau_r, \tau^{\star 1}_{po}, \tau_r^{\star 2}) = p^w_{eq}(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$. And, foreign country $\star 2$ prefers $\tau_{po}^{\star 2}$, pinned down by $W_{p^{\star 2}}^{\star 2} = 0$, given that it will be reciprocated by some $(\tau_r, \tau_r^{\star 1})$ such that $p_{eq}^w(\tau_r, \tau_r^{\star 1}, \tau_{po}^{\star 2}) = p_{eq}^w(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$. In general, each of these defines a locus on the iso-world-price surface in the $(\tau, \tau^{\star 1}, \tau^{\star 2})$ domain, which do not necessarily intersect each other at the same point - and if they do, which occurs when the political optimum world price coincides with the existing world price, I am in the knife-edge case where countries' desires are perfectly aligned. However, any deviation from this case implies that the home country is either on the long side or on the short side. And if the home country is either long or short, an externality situation between the two foreign countries arises under which each country would like to get its own preferred local price, while leaving all other requirements in an agreement to be fulfilled by other countries.

To describe the three cases in line with the previous section, note that the two foreign countries' preferred tariffs together imply a unique value for the home tariff, which I denote by τ_{ipo} satisfying $p_{eq}^w(\tau_{ipo}, \tau_{po}^{\star 1}, \tau_{po}^{\star 2}) = p_{eq}^w(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$. Thus, I again obtain the following cases: (1) If $\tau_{ipo} > \tau_{po}$, then home country is *long*: it prefers *more* liberalization than the foreign exporters jointly want. (2) If $\tau_{ipo} < \tau_{po}$, then home country is *short*: it prefers *less* liberalization than the foreign exporters jointly want. (3) If $\tau_{ipo} = \tau_{po}$, then home country is *equal*.

Whenever countries find themselves in either case (1) or (2), they have to bargain with each other over the allocation of concessions.[15] Depending on the economic environment, relative to the home country's desire, together the exporters may want too much from the home country in terms of tariff concessions (which would be the case, for example, when they have relatively larger market power), or too little from

[15]It is in case (2), where home is short and where therefore the two foreign exporters may face collective rationing, that the "rationing rule" used to allocate concessions across the two foreign exporters becomes relevant, and this is where my second departure from the assumptions maintained in Bagwell and Staiger (forthcoming) arises. As will become clear in the following sections, I allow the foreign exporters to bargain over the allocation of their concessions in this case, and this amounts to a rationing rule that is not allowed under the institutional assumptions maintained in Bagwell and Staiger (forthcoming).

the home country in terms of tariff concessions (which would be the case, for example, when they have relatively smaller market power). In order to maintain tractability while exploring this idea in a formal bargaining model, I will make some simplifying assumptions from now on, which will allow me to illuminate the interaction between institutions and bargaining outcome. This will also enable me to link the model with data in a more direct way. In particular, I will next move to a partial equilibrium representation of the three-country general equilibrium model I have just described. While this achieves the desired simplification, it also raises some conceptual issues which warrant discussion before proceeding further.

In theory, by relying on a partial equilibrium model, I lose a key feature of import tariffs that is exhibited in the general equilibrium model, namely, that due to Lerner Symmetry there is no formal distinction between an import tariff and an export tax/subsidy in terms of impacts on economic variables. Hence, in the general equilibrium model, reciprocity can be understood as a reciprocal exchange of tariff cuts by the negotiating countries, as I have previously described. And alternatively, in the general equilibrium model and due to Lerner Symmetry, reciprocity could be equivalently understood as a reciprocal exchange of, say, a home import tariff cut for foreign export tax cuts/export subsidy increases. With this in mind, note that Lerner Symmetry, however, does not apply in a partial equilibrium model. Consequently, in order to replicate the world-price-preserving impacts of reciprocity in a partial equilibrium model, I must have the exporters reciprocating the importer tariff cut with export policy changes of their own: exchange of import tariff cuts simply can not keep the world price fixed in a partial equilibrium setting.

In practice, exchange of tariff cuts for export tax cuts/export subsidy increases is counterfactual to GATT/WTO market access negotiations. However, this form of exchange should not be taken and interpreted literally, but should rather be seen as a simple way for a partial equilibrium model, simplifying but with more tractability, to reflect the Lerner-Symmetry-like features of the general equilibrium model previously sketched, while maintaining the key world-price-preserving feature of reciprocity that arises in a general equilibrium model, under the forms of reciprocal exchanges of tariff concessions that characterize the GATT/WTO market access negotiations. This treatment seems reasonable given that in GATT negotiations many tariffs are being negotiated and general equilibrium forces such as Lerner Symmetry are likely to be quite relevant in capturing the impacts of the round: that is, these negotiations are not restricted to a single narrow industry, where a focus on partial equilibrium forces would probably suffice, instead they cover many industries and represent broad changes in a country's tariff structure.

4. The Model

I now describe the 3-country partial equilibrium trade model within which I analyze the tariff bargaining problem. I first introduce a static environment, which translates into a dynamic setting once bargaining is included.[16] There are three countries and one good in this world, where a home country (with no superscript) is a natural importer of that good and two foreign ($\star 1$ and $\star 2$) countries are exporters of that same good. Regarding the protection instruments, the home country imposes import tax denoted as τ, and the foreign country $\star j$ $(j = 1, 2)$ imposes export subsidy denoted as $\tau^{\star j}$. Furthermore, the demand/supply system is denoted as $D(p)$, $Q(p)$ for the home country; $D^{\star j}(p^{\star j})$, $Q^{\star j}(p^{\star j})$ for foreign country $\star j$, where p is the domestic price of home country, $p^{\star j}$ is the local prices of foreign country j, and p^w is the world price. Throughout my analysis, I adopt the particular parameterization: $Q = 0$, $Q^{\star 1} = \omega$, $Q^{\star 2} = 1 - \omega$. $D = 1 - \alpha p$, $D^{\star 1} = \beta - p^{\star 1}$, $D^{\star 2} = \beta - p^{\star 2}$, that is,

[16]My model is similar to Bagwell and Staiger (2001)'s competing-exporter model. In terms of parametrization, it is also similar to Bagwell, Mavroidis and Staiger (2007) where a competing-importer, rather than competing-exporter, setup is used.

a 3-country endowment-economy model with linear demands.[17] Also, I define the home country's import demand as $M(p) \equiv D(p) - Q(p)$, and $\star j$'s export supply as $E^{\star j}(p^{\star j}) \equiv Q^{\star j}(p^{\star j}) - D^{\star j}(p^{\star j})$.

The prices are linked by the conditions $p(\tau, p^w) = p^w + \tau$, $p^{\star j}(\tau^{\star j}, p^w) = p^w + \tau^{\star j}$. Thus the market clearing condition $M(p(\tau, p^w)) = \sum_{j=1,2} E^{\star j}(p^{\star j}(\tau^{\star j}, p^w))$ yields the equilibrium world price $p^w = p^w(\tau, \tau^{\star 1}, \tau^{\star 2})$, given a tariff/subsidy combination $(\tau, \tau^{\star 1}, \tau^{\star 2})$. Following the literature on models with political economy, I assume each country's welfare function is a weighed sum of consumer surplus, producer surplus and tariff revenue (or subsidy cost). In particular, a politically motivated government will put higher weights on its producer surplus. I denote the set of political weights as $(\gamma, \gamma^{\star 1}, \gamma^{\star 2})$. Then the (static or stage) payoff functions in my partial equilibrium setting, with an extra political weight on producer surplus, can be written as:

$$W^k(\tau^k, \tau^{-k}) \equiv W^k(p^k, p^w) = \int_{p^k}^{\beta} D^k(p)dp + \gamma^k \int_{0}^{p^k} Q^k(p)dp + (p^k - p^w)[D^k(p^k) - Q^k(p^k)]$$

where $k \in \{\text{none}, \star 1, \star 2\}$. To avoid the (uninteresting) technical issue that countries may subsidize the exporter in a lump-sum fashion in my endowment-economy model, I assume that countries are restricted by the upper bound of local price $\bar{p}^{\star j} \equiv \beta$, that is, a policy that makes local price above $\bar{p}^{\star j}$, and hence eliminates any consumption distortion from further increases in the tariff, is not permitted.[18]

[17] This system enables me to consider the market power of the home country while keeping reciprocity in a simple form. In particular, as α becomes smaller, the home country's ability to manipulate the world price decreases, implying smaller market power. And, in terms of normalization, I am assuming the following conditions: foreign countries have identical demand function, the slope of which is equal to the intercept of the home country's demand function, and both of them are equal to the total supply. Under these conditions, the functions in my model can be interpreted as having been normalized by the total supply.

[18] Technically, this assumption ensures concavity in the payoff functions. In a sense, this is a restriction similar to Ossa (2011) where an (exogenous) upper bound on Nash policy is imposed. However, under more general demand/supply system, this restriction is not necessary.

Having described the trade framework, I now introduce the bargaining procedure.[19] The game consists of two periods.[20] Actions in the first period are trivial in the sense that countries set their (static) Nash policy choices. And, I assume countries are *bound* (restricted) in a way such that either they negotiate a new agreement under reciprocity, or they stick to their existing policies if they fail to negotiate a new one, which I interpret as an institutional constraint in my bargaining game reflecting the constraint imposed by previous GATT tariff bindings. Thus the policy profile in the first period will serve as the status quo in the second (bargaining) period. In order to determine the home country's strategy in a simple way, I assume that the home country makes a take-it-or-leave-it offer in the beginning of the second period. Then the foreign countries bargain with each other on how to make (reciprocal) concessions to the home country, given the home country's offer.[21] Also, in the beginning of the second period, there are possibilities that the foreign exporting countries may face their own (private) political shocks,[22] which they expect to last for indefinite amount

[19]While it seems that a timing game such as war of attrition, or (repeated) game of chicken, which is presumably simpler than a bargaining setting, is appropriate for modeling delay, I note several features that make a bargaining setting more appropriate in my tariff negotiation: (1) my actual tariff negotiation procedure takes the form of offers and counteroffers, which a timing game may not capture well. (2) Besides the conflict, reciprocity requires the countries to cooperate in some sense, instead of being 'winner takes all', i.e. stay-exit - it involves compromises between the players, thus the possible outcome is continuous, rather than binary. (3) Actual tariff negotiation involves variations in the time between offers and counteroffers, which I argue justifies the use of a framework with strategic delay.

[20]Essentially, what countries are doing is, starting from an initial Nash equilibrium, bargaining towards a new equilibrium, which they may have difficulties to determine in the shadow of potential political pressures. The period here can be interpreted as a "round" in the GATT/WTO framework.

[21]Note that this assumption is not as restrictive as it seems, since there isn't much intensive margin adjustment in the offers, as documented in Bagwell, Staiger and Yurukoglu (2017). Moreover, while this is a simplifying assumption, one justification (outside the current model) would be that the home country can switch partners costlessly. Then the home country might be thought as having the ability to make take-it-or-leave-it offers, as suggested by Fudenburg, Levine and Tirole (1987).

[22]While introducing the possibility of a political shock to the home country may conceptually make the model more complete, it will not contribute much to my analysis, as I assume that the home country only makes a take-it-or-leave-it offer. Moreover, in my later parametric setting, it does not matter whether the home country has political pressure or not, as I assume, for simplicity, that it does not have any domestic endowment in that setting.

of time.[23] Throughout the analysis, the institutional constraints I impose are the Reciprocity and MFN constraints. Under these two conditions, the world price should be kept at its existing level.[24] Thus the home country's offer on its (MFN) tariff τ^h has to be reciprocated by foreign countries' policy choices $(\tau^{\star 1}, \tau^{\star 2})$ in any agreement. Formally, I define the reciprocity constraint as follows.

Definition 1. Given an existing (MFN) policy profile $(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$, a negotiated (MFN) profile of tariff/subsidy $(\tau, \tau^{\star 1}, \tau^{\star 2})$ is said to satisfy *Reciprocity* if and only if $p^w(\tau, \tau^{\star 1}, \tau^{\star 2}) = p^w(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$.

To characterize the outcome when the countries fail to reach an agreement, I assume that the existing tariffs (status quo) are determined from the Nash equilibrium in the initial period, under the political pressure during that period. As the round gets under way, there is a new draw of political pressures in foreign countries.[25] For notational consistency, throughout the paper, I denote any existing variable, determined in the first period, with a subscript "0." In particular, recall that these variables are determined by the (static) Nash equilibrium in the first period.

I now proceed to the discussion of the bargaining instruments under reciprocity. Generally speaking, countries would have to decide on $(\tau, \tau^{\star 1}, \tau^{\star 2})$ in any agreement. However, since I endow the home country with the bargaining power to make a take-it-or-leave-it offer, τ will be simply fixed in any agreement, that is, the home country faces a static decision problem. Thus the policy profile $(\tau, \tau^{\star 1}, \tau^{\star 2})$ reduces to $(\tau^{\star 1}, \tau^{\star 2})$. Moreover, the reciprocity condition $p^w(\tau, \tau^{\star 1}, \tau^{\star 2}) = p^w(\tau_0, \tau_0^{\star 1}, \tau_0^{\star 2})$ enables

[23] This timing abstracts away some other subtleties when mapped into the data, which will be considered in more details that are consistent with the broad timing assumptions made here, in the empirical section.

[24] See Bagwell and Staiger (2002) for a full review of this result.

[25] In a sense, Jensen and Thursby (1990) analyzed a similar game where a new shock is generated in the second stage. While the shocks are correlated in their setting, I assume independence instead, so as to simplify my analysis such that existing tariff does not convey information about the countries' current political pressure.

me to rewrite $(\tau^{\star 1}, \tau^{\star 2})$ as $(\tau^{\star 1}, \tau^{\star 2}(\tau^{\star 1}))$. Based on this observation, an agreement could be simply characterized by a scalar, which I define, without loss of generality, as $g \equiv \tau^{\star 2} - \tau^{\star 1}$. Then any agreement $(\tau^{\star 1}, \tau^{\star 2})$ can be equivalently characterized by g given that they sum to a constant prescribed by the reciprocity. Thus I could simply assume that countries are bargaining on scalar g, which can be interpreted as the difference or *gap* between the two foreign countries' tariff/subsidy.[26] Alternatively, since $\tau^{\star j} = p^{\star j} - p^w$ and p^w is fixed at p^w_0 under reciprocity, I could equivalently characterize an agreement by $(p^{\star 1}, p^{\star 2})$, and assume countries are directly bargaining on their local prices, wherein the scalar could be well written as $g = p^{\star 1} - p^{\star 2}$. Thus, under this interpretation, countries are effectively bargaining on the difference or *gap* between the two foreign countries' local prices. In sum, given the home country's offer, an agreement/offer can be equivalently characterized by $(\tau^{\star 1}, \tau^{\star 2})$, $(p^{\star 1}, p^{\star 2})$ or g. That said, I will just rewrite countries' welfare in terms of g in the analysis of bargaining, and assume that countries are bargaining on g whenever it is convenient to do so, while keeping the above interpretation in the background.

After observing the initial take-it-or-leave-it offer from the home country, $\star 1$ and $\star 2$ will start the negotiation on how to coordinate/allocate their policy choices implied by reciprocity. This negotiation builds on the standard form of an alternating-offer bargaining. In this framework, $\star 1$ will propose first, then $\star 2$ will decide whether to accept the offer. If $\star 2$ accepts, then the game ends. Otherwise, it will be $\star 2$'s turn to make proposals and $\star 1$ responds. Moreover, I assume that each country can wait for as long as they want before making an offer or responding to an offer. Furthermore, each country's political parameter is private information. Thus, delay is a strategic device to communicate countries' types. And I impose the following assumptions and notations:[27]

[26] While Harstad (2007) links this scalar to *harmonization* of policies in an *ex ante* symmetric environment, the representation here is just for notational convenience in my setting.

[27] My assumptions are similar to those adopted in Admati and Perry (1987) and Harstad (2007).

(A1) $\gamma^{\star j} \in \{l, h\}$, $p \equiv Pr(\gamma^{\star j} = h) = \frac{1}{2}$, where$\gamma^{\star j}$ is private information and $h > l = 1$.[28]

Thus, I assume that as a result of the shock, in the second period (and forever thereafter) each foreign country is either a national income maximizer, and so its politically optimal reaction curve tariff corresponds to free trade, or it has some political pressure.[29]

(A2) A common discount factor $\delta = e^{-\rho}$ for the two foreign countries, where $\rho > 0$ is an instantaneous discounting rate, and the length of the unit time between offers is arbitrarily small.[30]

I denote $\star 1$'s and $\star 2$'s instantaneous welfare function simply as $W^{\star 1}(g, \gamma^{\star 1})$ and $W^{\star 2}(g, \gamma^{\star 2})$ respectively, for an agreement g,[31] and their inside option (status quo) as $W_0^{\star 1}(\gamma^{\star 1})$ and $W_0^{\star 2}(\gamma^{\star 2})$. Thus the total (discounted) welfare that a foreign country obtains from an agreement g, reached at time t_A is:

$$\int_0^{t_A} \delta^s W_0^{\star j}(\gamma^{\star j}) ds + \int_{t_A}^{\infty} \delta^s W^{\star j}(g, \gamma^{\star j}) ds = \frac{1}{\ln(\delta^{-1})} \{\delta^{t_A}[W^{\star j}(g, \gamma^{\star j}) - W_0^{\star j}(\gamma^{\star j})] + W_0^{\star j}(\gamma^{\star j})\}$$

In the special case with no delay, i.e. $t_A = 0$, this welfare becomes $\frac{1}{\ln(\delta^{-1})} W^{\star j}(g, \gamma^{\star j})$. With delay, i.e. $t_A > 0$, the welfare will diminish according to the $\frac{1}{\ln(\delta^{-1})}\delta^{t_A}$ term. To allow for both the case when home is *short* and the case when home is *long*, and to preserve the trade pattern in the initial equilibrium, namely (1) positive export volume by $\star 1$ and $\star 2$, (2) positive domestic consumption in $\star 1$ and $\star 2$ and (3) positive endowment, I impose the following restrictions on the model parameters:

[28] In the literature on bargaining with private information, the space of types are usually limited. With a continuum of types, it is generally known that a great many of equilibria exist, undermining the model's predictive power. Cramton (1992) considers such a framework. And I could alternatively interpret the political shock as the *possibility* of a new draw, which seems more natural in this setting.

[29] This may yield a policy (choice) set similar to Maggi and Staiger (2011) in *equilibrium*, which consists of Free Trade and Protection. However, my restriction is directly on the possible set of parameters characterizing political pressure.

[30] Note that I do not restrict the discount factor of the home country.

[31] Recall that I have previously established the scalar representation of an agreement.

(A3) $\beta \in [\frac{1}{2}, \frac{3}{2})$, $(2\beta - 1)\alpha \leq 2$ and $\omega \in \Omega(\alpha, \beta)$.[32]

To formally establish the bargaining setting, I introduce the following notation for this bargaining game under incomplete information. A history H^n after n offers, consisting of the series of offers and counteroffers, is recursively defined as $H^0 \equiv \{(\emptyset, 0)\}$ and $H^n \equiv H^{n-1} \cup \{(g^n, t^n)\}$ for $n \geq 1$, where g^n is the offer made at t^n. Note that whenever an offer is accepted, the game ends. The action space is defined as $\mathcal{A}^n = \mathcal{A} \equiv \{\text{accept}\} \cup (\mathbb{G}, \mathbb{R}^+)$ for $n \geq 1$ and $\mathcal{A}^0 \equiv (\mathbb{G}, \mathbb{R}^+)$, where $\mathbb{G}$ is the *possible* range of agreements. A pure strategy for a country $\star j$, when it is its turn to propose/respond, is defined as a mapping from the set of history to the action space. In particular, I assume when n is even, it is $\star 1$'s turn to propose/respond, thus $\star 1$'s strategy can be written as $\{s_1^{2k}\}_{k \geq 0}$, with $s_1^{2k} : H^{2k} \mapsto \mathcal{A}$ for $k \geq 1$ and $s_1^0 : H^0 \mapsto \mathcal{A}^0$. Otherwise it is $\star 2$'s turn to take actions, thus $\star 2$'s strategy can be written as$\{s_2^{2k+1}\}_{k \geq 0}$ with $s_2^{2k+1} : H^{2k+1} \mapsto \mathcal{A}$.[33] Thus, when responding to an offer g^n made at t^n, a country can accept it, or respond with a counter offer g^{n+1} at $t^{n+1} \geq t^n$. Moreover, define $\mathcal{H} \equiv \{H^n\}_{n \geq 0}$. Let $\mu^j : \mathcal{H} \mapsto [0, 1]$ be $\star j$'s belief on the probability of $\star - j$ being of high political pressure, i.e. $\gamma^{\star - j} = h$. In particular, I assume $\mu^1(H^0) = \mu^2(H^0) = p \equiv \frac{1}{2}$.

Due to presence of private information in this game, I will focus on its sequential equilibrium, which requires the strategy profile $\{(s_1^{2k}, s_2^{2k+1})\}_{k \geq 0}$ to be optimal after every history, given the beliefs (μ^1, μ^2), and the beliefs have to be consistent with the strategy profile in the sense of Bayesian updating. Moreover, to ensure uniqueness of the equilibrium, I impose the intuitive criterion as my refinement on the off-equilibrium beliefs.

[32] See the Appendix for the definition of $\Omega(\alpha, \beta)$, and the formal rationale and derivation underlying these restrictions. For example, as an application in the next section where it is assumed that $\alpha = 1$ and $\beta = \frac{1}{2}$, I will need to restrict $\omega \in \Omega(\alpha, \beta) = [\frac{1}{6}, \frac{5}{6}]$.

[33] As will be seen later, the proposing order matters, in the sense that it will affect the expected delay of the game, as the countries in my model is *ex ante* asymmetric.

5. Analysis

As a benchmark set of trade policies, the Nash Equilibrium (countries choosing their optimal tariff/subsidy) can be characterized by the best response equations $\{\frac{\partial W^k(\tau^k,\tau^{-k})}{\partial \tau^k}=0\}_{k\in\{\text{none},\star 1,\star 2\}}$:

$$\{-(1-\gamma^k)Q^k(p^k)+(p^k-p^w)[D'^k(p^k)-Q'^k(p^k)]\}\cdot(1+\frac{\partial p^w}{\partial \tau^k})-[D^k(p^k)-Q^k(p^k)]\frac{\partial p^w}{\partial \tau^k}=0$$

Thus, under Nash Equilibrium, countries are not only motivated by their local prices, which represents their political pressure, but also concerned with the terms-of-trade effect, characterized by $\frac{\partial p^w}{\partial \tau^k}$. As a result, countries have the incentive to manipulate their terms-of-trade by shifting the associated cost to others, as long as they have some market power, i.e. as long as they are not "small" in international markets. In particular, as shown in the Appendix, I have $E_0^{\star 1} > E_0^{\star 2} \Leftrightarrow \omega > \frac{1}{2}$ in the equilibrium of the first period. In other words, the country with a larger endowment will be the *principal supplier*, referring to the foreign country that is the largest export supplier to the home country of the good under consideration. As will be seen in the analysis of the bargaining equilibrium, this allows me to examine the effect of another institutional rule in GATT/WTO, in addition to MFN and reciprocity in negotiations, namely the Principal Supplier rule which leads countries to negotiate concessions with a principal supplier on a given product. In my bargaining model, I will capture this institutional feature by allowing the principal supplier to make the first offer, and I will compare the delay under this rule to the alternative in which the smaller supplier makes the first offer.[34]

[34]Bagwell and Staiger (forthcoming) suggest a different institutional interpretation of the principal supplier rule, under which when foreign exporters are collectively rationed the principal supplier assumes the "first-priority position" and receives the trade volume offered by the home country up to its preferred level, with the second exporter then assuming the role of the "residual claimant". Under this interpretation of the principal supplier rule, there is no bargaining between the foreign suppliers and hence no reason for strategic behavior (such as delay). When viewed together, my paper and Bagwell and Staiger (forthcoming) suggest the importance of understanding exactly how

As discussed in the context of my general equilibrium trade models, the politically optimal reaction curve tariff is the optimal tariff when a government does not value the terms-of-trade effect when making its policy choice, which solves $W_p = 0$ for any arbitrarily given p^w.[35] Also recall that the motive of each country to achieve this tariff through tariff bargaining is the source of the externality across the competing exporters in my model. Throughout the paper, I will simply refer to each country's politically optimal reaction curve tariff as its "bliss point tariff", or simply its bliss point. Alternatively, I can interpret the bliss point tariffs as those that emerge when countries behave *as if* they were small countries and set their tariff/subsidy to attain the preferred local prices.[36] Concretely, this set of bliss points can be formally characterized by $\{\frac{\partial W^k(p^k,p^w)}{\partial p^k} = 0\}_{k\in\{\text{none},\star 1,\star 2\}}$:

$$-(1-\gamma^k)Q^k(p^k) + (p^k - p^w)[D'^k(p^k) - Q'^k(p^k)] = 0$$

In particular, the bliss points in my model are $\bar{\tau}_{po}^{\star 1} = (\gamma^{\star 1}-1)\omega$, $\bar{\tau}_{po}^{\star 2} = (\gamma^{\star 2}-1)(1-\omega)$, $\tau_{po} = 0$, $p_{po}^w = \frac{2\beta+(1-\gamma^{\star 1})\omega+(1-\gamma^{\star 2})(1-\omega)}{2+\alpha}$. Given the restriction $p^{\star j} \leq \bar{p}^{\star j}$ which implies $\tau^{\star j} \leq \bar{\tau}^{\star} \equiv \frac{(2\beta+1)\alpha-1}{2+4\alpha}$, there is a technical issue in the current setting, namely, the above bliss points may not be attainable if $\bar{\tau}_{po}^{\star j} > \bar{\tau}^{\star}$.[37] To handle this problem, I now define the (constrained) bliss point policies as $\tau_{po}^{\star 1} \equiv \min(\bar{\tau}^{\star}, \bar{\tau}_{po}^{\star 1})$, $\tau_{po}^{\star 2} \equiv \min(\bar{\tau}^{\star}, \bar{\tau}_{po}^{\star 2})$,

the principal supplier rule works in practice to shape GATT tariff bargaining, something that future analysis of the GATT bargaining data may make possible.

[35]The politically optimal reaction curve tariff is originally defined in Bagwell and Staiger (1999, 2002). Note that when all countries are on their politically optimal reaction curves, the p^w is in fact the one that would be implied by all the politically optimal reaction curve tariffs simultaneously, which Bagwell and Staiger define as *political optimum*.

[36]While each of these two interpretations is appropriate under perfect competition, they may not be equivalent under imperfect competition. See Bagwell and Staiger (2012) for an example.

[37]Specifically, countries' payoff functions will be *linear* when the local price exceeds the choke price above which domestic demand is driven to zero, creating a kink in the payoff function. Thus the (interior) bliss point previously characterized by a first order condition becomes invalid whenever it results in a local price larger than the choke price. However, this issue will not be present under a more general demand/supply system.

which specifies that the bliss point $\tau^{\star j}_{po}$ will be set to $\bar{\tau}^{\star}$ whenever $\bar{\tau}^{\star j}_{po} > \bar{\tau}^{\star}$. And, unless explicitly expressed, I will refer to $(\tau^{\star 1}_{po}, \tau^{\star 2}_{po})$ as bliss points from now on.

The constraint of reciprocity requires that $p^w = p^w_0$ in any new agreement, implying $\tau^{\star 1} + \tau^{\star 2} = \frac{-[(2\beta-1)\alpha-2](\alpha-1)}{2+4\alpha} - \alpha\tau \equiv \pi$. Thus, the concession jointly preferred by $\star 1$ and $\star 2$, $\tau^{\star 1}_{po} + \tau^{\star 2}_{po}$ can be less than, equal to or larger than π. As a result, again, there are generally three possibilities for a given pair of political pressure.[38] Moreover, in the Appendix, I show that the home country will make the offer $\tau = 0$ provided that it is sufficiently patient and the parameter restriction $\alpha \leq \min(\frac{\sqrt{17}-1}{2}, \frac{2}{2\beta-1})$ holds. In order to avoid the complication associated with the home country's strategy, I henceforth assume that the home country is also sufficiently patient, $\frac{2}{2\beta-1} > \frac{\sqrt{17}-1}{2}$ and $\alpha \leq \frac{\sqrt{17}-1}{2}$, which ensures that the home country's offer is simply $\tau = 0$. Next I formally discuss the cases in terms of the home country's "position" (*short, long* or *equal*). Note that $\pi = \frac{-[(2\beta-1)\alpha-2](\alpha-1)}{2+4\alpha}$ under $\tau = 0$. Since $(2\beta - 1)\alpha < 2$, I can discuss the following possibilities regarding the size of the reciprocal concession asked by the home country, based on the parameters:

(1) (*home short*) If $\alpha < 1$, that is the home country has less market power, then there is too little concession ($\pi < 0$) to be jointly made by $\star 1$ and $\star 2$ in order to reciprocate the home country. Thus $\star 1$ and $\star 2$ will compete to make more concessions. In other words, the home country's offer requires at least one of the two foreign countries to levy export tax in an agreement, although neither of them prefers to do so.

(2) (*home long*) If $\alpha > 1$, that is the home country has more market power, then a l-type $\star 1$ and a l-type $\star 2$ will jointly want less concession than that asked by the home country ($\pi > 0$), implying that they will compete to make less concession. And, as can be seen from Lemma 4 in the Appendix, if the size of h-type's political pressure is small enough, then they compete to make

[38] Bagwell and Staiger (forthcoming) also analyzes a framework with these possibilities.

less concession in the entire game. In other words, when the home country asks for too much liberalization, $\star 1$ and $\star 2$ has to allocate the extra (costly) liberalization relative to their bliss points. However, for given (α, β, ω), if the size of h-type's political pressure, h, is large enough, then they will compete to make more concessions.

(3) (*home equal*) If $\alpha = 1$, the concession asked by the home country is just enough ($\pi = 0$) for a l-type $\star 1$ and a l-type $\star 2$, that is, free trade will be the best for every country. However, if either of the two countries has any political pressure, that is, being a h-type, then they will compete to make more concessions. Thus, in this case, $\star 1$ and $\star 2$ weakly compete to make more concessions.

Also recall that the parameter restrictions that I have imposed so far is $\beta \in [\frac{1}{2}, \frac{\sqrt{17}+5}{8})$, $\alpha \leq \frac{\sqrt{17}-1}{2}$ and $\omega \in \Omega(\alpha, \beta)$. In particular, for $\alpha \leqslant 1$, the home country will be (weakly) on the short side, that is, foreign countries compete for making (weakly) more concessions. I analyze a simple version of this case next and leave the discussion of the home long case to the Appendix.[39]

For simplicity, I impose $\alpha = 1$ and $\beta = \frac{1}{2}$ in this subsection to capture the case of *home short*.[40] In line with assumption **(A3)**, I focus on $\omega \in \Omega(1, \frac{1}{2}) = (\frac{1}{6}, \frac{5}{6})$. For convenience, I denote $g_0^{\star 1}$ and $g_0^{\star 2}$ as the two foreign countries' status quo policy, respectively. Then, $\star 1$'s welfare function and $\star 2$'s welfare function can be conveniently

[39] The case of home short is also of independent interest in its own right. Recall that relative to Bagwell and Staiger (forthcoming) I have relaxed two institutional assumptions: first, the restrictions implied by reciprocity in renegotiation; and second the restrictions implied by the class of rationing rules considered by Bagwell and Staiger (forthcoming). Maintaining the first restriction while relaxing only the second would rule out the home long case, leaving the home short (and the knife-edge home equal) case as the only relevant case to consider.

[40] Note that the results in what follows stand as long as the home country is short, but not too short. In particular, when home is sufficiently short, the model reduces to one that is very similar to Harstad (2007) where monotone payoff functions are present, with non-monotonicity in the relationship between delay and externality. And I will return to this qualification when examining the evidence from the GATT bargaining data.

represented in terms of g as

$$W^{\star 1}(g,\gamma^{\star 1}) = \frac{1}{8} + p_0^w(\omega - \frac{1}{2}) + [(\gamma^{\star 1}-1)\omega + p_0^w](p_0^w - \frac{g}{2}) - \frac{1}{2}(p_0^w - \frac{g}{2})^2$$

$$W^{\star 2}(g,\gamma^{\star 2}) = \frac{1}{8} + p_0^w(\frac{1}{2} - \omega) + [(\gamma^{\star 2}-1)(1-\omega) + p_0^w](p_0^w + \frac{g}{2}) - \frac{1}{2}(p_0^w + \frac{g}{2})^2$$

And their payoffs from existing policies are $W_0^{\star 1}(\gamma^{\star 1}) = W^{\star 1}(g_0^{\star 1},\gamma^{\star 1})$ and $W_0^{\star 2}(\gamma^{\star 2}) = W^{\star 2}(g_0^{\star 2},\gamma^{\star 2})$ respectively. Starting from their existing trade policies, each country would like higher local prices, until their bliss points are reached, that is, $\frac{\partial W^{\star 1}}{\partial g}|_{g=g_0^{\star 1}} < 0$ and $\frac{\partial W^{\star 2}}{\partial g}|_{g=g_0^{\star 2}} > 0$. Also, I can write the bliss point of a type-i foreign country k as $g_{po}^{\star k}(i) \equiv \underset{g}{\text{argmax}}[W^{\star k}(g,i) - W_0^{\star k}(i)]$. As is presented in the Appendix, it can be shown that $g_{ij} \in (g_{po}^{\star 1}(i), g_{po}^{\star 2}(j))$, that is, the agreement has to lie in between the two foreign countries' bliss points. Moreover, the possible range for any offer g is $\mathbb{G} \equiv [g_0^{\star 2}, g_0^{\star 1}]$, as any offer outside of this range will be certainly rejected by one or both of the foreign countries.

Recall that, although in reality each exporter is negotiating bilaterally with the importer, I am formally looking at this equivalently as the two exporters making offers to each other rather than to the importer about how they are going to jointly reciprocate the importer's tariff cut offer. In the case of home short, whenever a country makes an offer that is attractive to a country of l type but unattractive to a country of h-type, then the other country will believe it is of l type. In particular, countries have the incentive to mimic a certain type, according to their own advantage. Specifically, when the home country is *short* as is analyzed in my current setting, being a type h is relatively advantageous (l may pretend to be h, but h never pretends to be l), as foreign countries compete to make concessions.[41] Let t_{ij} be the equilibrium

[41]This is because a high pressure country will always want larger share of concession in the range of feasible agreements. Technically, it is because $g_{l\cdot} < g_{h\cdot} < g_{po}^{\star 1}(h)$ and $g_{\cdot l} < g_{\cdot h} < g_{po}^{\star 2}(h)$. Intuitively,

amount of delay, or time to agreement, when $\star 1$ is of type i and $\star 2$ is of type j, where $i, j \in \{h, l\}$. In particular, when both countries are of high type, then $\star 1$ delays its offer until t_{hl}, and $\star 2$ delays further by $t_{hh} - t_{hl}$, making its counteroffer at t_{hh}. Thus a bargaining outcome can be represented as (g_{ij}, t_{ij}), which specifies the agreement (g_{ij}) and the time to agreement, that is, delay (t_{ij}). As I will show later on, the equilibrium agreement (g_{ij}) is one that coincides with the complete information version of this game. Thus, I summarize the outcome in the following table, where the row labels represent the first proposer ($\star 1$)'s political pressure, while the column labels represent the second proposer ($\star 2$)'s political pressure.

type	h	l
h	(g_{hh}, t_{hh})	(g_{hl}, t_{hl})
l	(g_{lh}, t_{lh})	(g_{ll}, t_{ll})

In a separating equilibrium, a type h may need to use delay as a (costly) device to truthfully reveal its type, to convince the other country that it is not of l type. While a type l does not need to signal its type by delay. Thus the incentive compatibility conditions only need to be binding so as to prevent l from mimicking h. The intuitive criterion ensuring that the delay will be just enough to signal the types, which pins down the unique equilibrium in terms of final agreement (g, t). Moreover, since ω is the parameter of interest, I use $g_{ij}(\omega)$ with $i, j \in \{l, h\}$ to highlight the dependency of g_{ij} on ω. Formally, I have the following lemma establishing the equilibrium.

Lemma 1. *The equilibrium agreement g coincides with that of the complete information counterpart, characterized by the standard Nash bargaining solution, and the*

high political pressure in the export country in my model can be thought of as the exporting country government facing high political pressure from its exporters who want *more* trade volume, relative to its import-competing producers who want *less* trade volume, which then translates into the high-pressure exporting country wanting to do a greater share of the reciprocating of the importing country's tariff-cut offer in order to capture a greater share of the increased market access for its exporters.

associated bargaining delay is characterized by the following components: $t_{ll} = 0$, $t_{lh} = 0$, *and*

(5.1)
$$\delta^{t_{hh}-t_{hl}} = \min\{\frac{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{hl}}{2})[-\frac{g_{hl}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{hh}}{2})[-\frac{g_{hh}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}, 1\}$$

(5.2)
$$\delta^{t_{hl}} = \min\{\frac{(\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{lh}}{2})[\frac{g_{lh}}{4} + \frac{1}{6}\omega - \frac{1}{36}] + (\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{ll}}{2})[\frac{g_{ll}}{4} + \frac{1}{6}\omega - \frac{1}{36}]}{\delta^{t_{hh}-t_{hl}}(\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{hh}}{2})[\frac{g_{hh}}{4} + \frac{1}{6}\omega - \frac{1}{36}] + (\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{hl}}{2})[\frac{g_{hl}}{4} + \frac{1}{6}\omega - \frac{1}{36}]}, 1\}$$

.

Given this lemma, without causing confusion, I hence denote $g_{ij}(\omega)$, or simply g_{ij} both as the Nash bargaining solution and as the equilibrium agreement in my model. Now I proceed to the equilibrium analysis for the remainder of this section and discuss several properties of the equilibrium that are relevant for the GATT/WTO institution. Moreover, these properties will also guide my empirical work later on.

Corollary 1. *Under the given parametric setting, whenever the first proposer has no political pressure, i.e. a national income maximizer, there is no delay in reaching an agreement.*

In other words, a necessary condition for the existence of delay is that the first proposer has some political pressure. As the relationship between delay and the market shares of the two exporters depends not only *directly* on ω, but also *indirectly* on $g_{ij}(\omega)$, the overall effect of ω on delay is not obvious. On the one hand, when ω changes, countries' payoffs are altered even for a fixed agreement. On the other hand, the agreement itself responds to ω, thus creating a second type of effect. I begin with the analysis of how $g_{ij}(\omega)$, which coincides with the Nash bargaining solution, depends on the market share.

Lemma 2. *As the first proposer's market share increases, its equilibrium local price increases, that is, export subsidy (export tax) strictly increases (decreases), whenever* *at least one of the two foreign countries are of high political pressure:* $\frac{\partial g_{lh}}{\partial \omega} > 0$, $\frac{\partial g_{hl}}{\partial \omega} > 0$, $\frac{\partial g_{hh}}{\partial \omega} > 0$. *Moreover, the country with higher political pressure always ends up with an export subsidy in equilibrium, while the country with lower political pressure ends up with an export tax, that is,* $g_{hl} < 0$, $g_{lh} > 0$. *When both of the countries are of high political pressure, the smaller (larger) exporter imposes an export subsidy (export tax), that is,* $sign(g_{hh}) = sign(\omega - \frac{1}{2})$.

When both of the countries are of high type, then the big supplier ends up with an export tax, while the small supplier ends up with an export subsidy, which seems to suggest the small supplier being a winner. However, note that the sign of g's depends not only on their bliss points but also their status quo. While the big supplier takes home with an export tax, it does so from the status quo with relatively large export tax, which is very bad when the political shock comes. Thus in this regard, the small supplier need not to be a winner.

I now proceed to the analysis of the relationship between delay and the export market characteristics. First, I focus on the delay with which a high-type $\star 2$ responds to $\star 1$'s initial offer, which I denote by $D_h^{\star 2}$. $D_h^{\star 2}$ is an important component of the total delay to agreement, and it is also the analytically easiest component to consider. Note that $D_{hh}^{\star 2} \equiv t_{hh} - t_{hl}$ as given by equation (5.1) in Lemma 1.

Note that if $\omega = \frac{1}{2}$ that is, symmetric suppliers, then I must have $D_h^{\star 2} > 0$, because $-g_{hl} > g_{hh} = 0$. In words, this component of delay is strictly positive at least when countries' market shares are similar. And, starting from $\omega = \frac{1}{2}$, $D_h^{\star 2}$ decreases with respect to ω to the point where $-g_{hl}(\omega) = g_{hh}(\omega)$ such that $D_h^{\star 2} = 0$. I denote this ω-cutoff as $\bar{\omega}$. And for $\omega \in [\frac{1}{2}, \bar{\omega})$, that is, $D_h^{\star 2} > 0$, the expression for $\frac{d(D_h^{\star 2})}{d\omega}$ is:

$$\frac{1}{\ln(\delta)}[\underbrace{\frac{-\frac{g_{hl}}{4}}{(\frac{5}{18}-\frac{1}{3}\omega+\frac{g_{hl}}{2})(-\frac{g_{hl}}{4}+\frac{5}{36}-\frac{1}{6}\omega)}\frac{\partial g_{hl}}{\partial\omega}-\frac{-\frac{g_{hh}}{4}}{(\frac{5}{18}-\frac{1}{3}\omega+\frac{g_{hh}}{2})(-\frac{g_{hh}}{4}+\frac{5}{36}-\frac{1}{6}\omega)}\frac{\partial g_{hh}}{\partial\omega}}_{\oplus}+$$

$$\underbrace{(\frac{1}{(\frac{5}{18}-\frac{1}{3}\omega+\frac{g_{hl}}{2})(-\frac{g_{hl}}{4}+\frac{5}{36}-\frac{1}{6}\omega)}-\frac{1}{(\frac{5}{18}-\frac{1}{3}\omega+\frac{g_{hh}}{2})(-\frac{g_{hh}}{4}+\frac{5}{36}-\frac{1}{6}\omega)})(\frac{1}{9}\omega-\frac{5}{54})}_{\ominus}]$$

On the one hand, for fixed proposals (g_{hl} and g_{hh}), as ω increases,[42] the direct effect, captured by $\ominus$, is that a l-type $\star 2$'s gain from the agreement, whether mimicking or not, decreases since it becomes a smaller supplier, that is, it benefits less for having smaller market share: in the limit when $\omega = \frac{5}{6}$, $\star 2$ is simply driven to autarky under the initial Nash equilibrium - in this case, its existing (Nash) policy, which is free trade, coincides with its bliss point, which means that a l-type $\star 2$ cannot gain anything from an agreement with $\star 1$. That said, whether $\star 2$ mimics or not, its payoff decreases with a larger ω. However, this reduction in benefit for $\star 2$, being a smaller supplier, is asymmetric between mimicking and non-mimicking. Concretely, the reduction when mimicking a h-type (thus getting an agreement g_{hh}) is *smaller* than that of sticking to its true type (thus an agreement g_{hl}). As a result, it provides *more* incentive for a l-type to mimic a h-type, which implies *longer* delay is needed to signal a h-type. In other words, when ω increases, the difference between mimicking and non-mimicking becomes larger. On the other hand, as ω increases, the indirect effect, captured by $\oplus$, through (g_{hl}, g_{hh}) is such that g_{hl} becomes more attractive to a l-type, while g_{hh} becomes less attractive. This provides *less* incentive for a l-type to mimic a h-type, which implies that *shorter* delay is needed to signal a h-type. Given these two competing forces, it is helpful to highlight them by a graph.

I use Figure 5.1 to illustrate the effect of ω on delay. In this figure, I focus on the delay *after* $\star 1$ has been revealed to be of type h, with an outstanding offer g_{hl} on the

[42]As previously noted, there is a one-to-one mapping between ω and export volume.

table. Given this proposal, $\star 2$ strategically delays its counteroffer, if any, to signal its type. In particular, a l-type $\star 2$ accepts g_{hl} without delay, while a h-type makes an counteroffer g_{hh} after delaying sufficiently long. Concretely, I plot the proposal (g) on the vertical axis against delay (t) along the horizontal axis, where the three colored loci represent l-type $\star 2$'s indifference curves, facing the tradeoff between cost of delay and gain from a better deal. As is standard in signaling games, the amount of delay (equilibrium under intuitive criterion) is pinned down when the l type becomes indifferent between accepting g_{hl} with no delay, and accepting g_{hh} with the equilibrium delay. This figure shows how ω affects the equilibrium outcome, where I decompose the transition from an initial equilibrium point A (it depicts the equilibrium when $\omega = 0.5$, which is where l-type $\star 2$'s indifference curve, passing g_{hl}, intersects the g_{hh} line) to a new equilibrium point D after ω increases. I can decompose this transition, denoted as $\overrightarrow{AD}$, to several components as follows. First, when ω increases, for a fixed proposal g_{hl}, the indifference curve shifts to the right, causing delay to increase to point B, and I label this 'direct effect' as $\overrightarrow{AB}$. Intuitively, this is due to the fact that as ω increases, delay becomes less costly for l-type$\star 2$ since it gains less from any given agreement, which I summarizes in the following lemma.

Lemma 3. *When ω increases, the difference between the small supplier ($\star 2$)'s existing Nash policy and its preferred policy as l-type becomes smaller, and vanishes in the limit. As a result, it tends to become indifferent between its existing policy and an agreement under reciprocity.*

In particular, as it becomes a smaller supplier, $\star 2$ cares less about agreement as its existing policy converges to its bliss point, willing to endure more delay for a given increase in its local price toward its bliss price. Second, larger ω leads to an increase in g_{hl}, which shifts the indifference curve from point B to point C, and I may name this 'indirect effect' as $\overrightarrow{BC}$, causing the reduction of delay. This is due to that

$\star 2$'s bargaining power increases as it becomes smaller, resulting in a better proposal offered by $\star 1$, thus shifting $\star 2$'s indifference curve inward. Finally, another indirect effect (denoted as $\overrightarrow{CD}$) is that larger ω shifts up the g_{hh} line, whose intersection with $\star 2$'s indifference curve determines the equilibrium outcome. The direction of this effect is similar to that of the $\overrightarrow{BC}$ effect, due to that the peak of indifference curve is at $g = 0$ for a national income maximizer.

Based on the argument above, I could write the total effect $\overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD}$. That said, whether delay increases or decreases with respect to ω depends which of the three effects dominate the others. A general pattern is that the direct effect ($\overrightarrow{AB}$) increases delay, while the indirect effects ($\overrightarrow{BC}$ and $\overrightarrow{CD}$) reduces delay. As established in Proposition 1, the indirect effects dominate the direct effect.[43]

Proposition 1. *For $\omega \in [\frac{1}{2}, \bar{\omega}]$, when both of the countries are of high political pressure, then the delay from the first offer to the second (final) offer is decreasing with respect to ω. For $\omega \in [\bar{\omega}, \frac{5}{6})$, there is no delay from the first offer to the second offer, where the agreement itself serves as a signaling device.*

This result links the delay (conditional on both countries are of high political pressure) to the export market shares, or the export market concentration. In general, the more concentrated the export market is, the shorter the delay from the first offer to the second offer is. Concretely, if I introduce the Herfindahl-Hirschman Index (HHI), defined as the sum of squared market shares, as a measure of the export market concentration in my model, that is, $HHI(\omega) \equiv (\frac{E_0^{\star 1}}{E_0^{\star 1}+E_0^{\star 2}})^2 + (\frac{E_0^{\star 2}}{E_0^{\star 1}+E_0^{\star 2}})^2 = \frac{9}{2}(\omega - \frac{1}{2})^2 + \frac{1}{2}$, then the market concentration is monotonically increasing in $\omega \in [\frac{1}{2}, \frac{5}{6})$. Thus the

[43]However, if the assumption $l = 1$ is relaxed, then as long as l becomes sufficiently large (thus $h > l$ should be large too), the direct effect dominates the indirect effects, which causes a positive association between ω and delay. Intuitively, when both l and h types have very high pressure, being a high type or low type does not matter much for the bargaining in the sense that all proposals tend to converge to a constant $g = \frac{2}{3}\omega - \frac{1}{3}$. This implies that there is infinitely small difference among g_{ll}, g_{lh}, g_{hl} and g_{hh}, rendering tiny indirect effects.

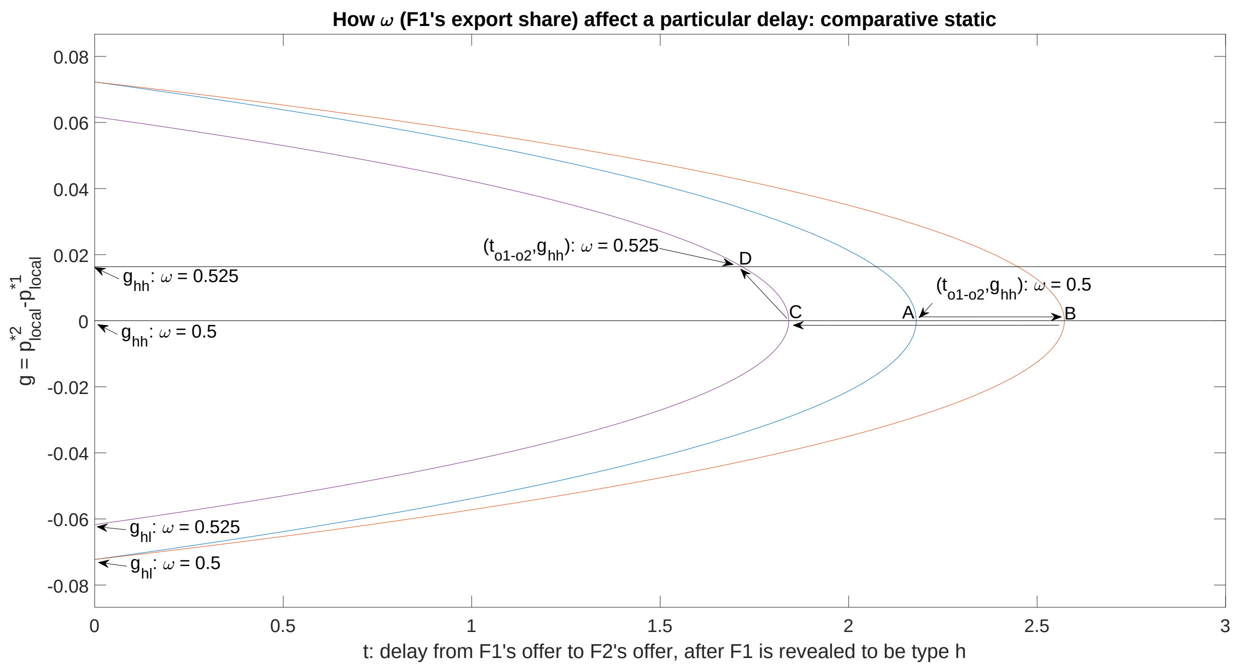

FIGURE 5.1. An illustration of ω's effect on delay

proposition above predicts a negative relationship between this particular delay and HHI.

Next, I turn to analyze the delay with which a high-type $\star 1$ makes the first offer, which I denote by $D_h^{\star 1}$. Note that $D_h^{\star 1} \equiv t_{hl}$ as given by equation (5.2) in Lemma 1. Starting from $\omega = \frac{1}{2}$, $D_h^{\star 1}$ stays at 0 until $\hat{\omega}$ where $(\frac{1}{3}\hat{\omega} - \frac{1}{18})^2 - (\frac{g_{lh}}{2})^2 = \delta^{t_{hh}-t_{hl}}(\frac{1}{3}\hat{\omega} - \frac{1}{18})^2 - (\frac{g_{hh}}{2})^2 - (\frac{g_{hl}}{2})^2$.

Proposition 2. *If the first proposer is of high political pressure, then for $\omega \in [\frac{1}{2}, \hat{\omega}]$ there is no delay from the beginning to the first offer. For $\omega \in (\hat{\omega}, \frac{5}{6})$, there exists delay from the beginning to the first offer, which is increasing then decreasing in ω (HHI).*

As can be seen from equation (5.2) in Lemma 1, other things equal, $D_h^{\star 1}$ and $D_h^{\star 2}$ are negatively correlated with each other. Intuitively, when $D_h^{\star 2}$ becomes shorter, a high-type $\star 1$'s incentive to misrepresent itself increases and this leads to $D_h^{\star 1}$ increasing in

ω initially. However, as ω becomes large enough such that $D_h^{\star 2} = 0$, the change in the equilibrium agreements will dominate, and similar to the arguments in analyzing $D_h^{\star 2}$, this drives a negative relationship between $D_h^{\star 1}$ and ω.

I summarize the findings in the following corollary.

Corollary 2. *(a) For $\omega \in [\frac{1}{2}, \hat{\omega}]$, there exists positive bargaining delay iff both of the countries are of high political pressure, which is the time elapsed from the first offer, made in the beginning, to the second offer. (b) For $\omega \in [\hat{\omega}, \bar{\omega}]$, there exists positive bargaining delay iff the first proposer is of high political pressure, which is the time elapsed from the beginning to the first offer plus that from the first offer to the second offer. (c) For $\omega \in [\bar{\omega}, \frac{5}{6}]$, there exists positive bargaining delay iff the first proposer is of high political pressure, which is the time elapsed from the beginning to the first offer, where the second offer is made and accepted immediately.*

After determining the "components" of delay and their relationship with ω, the relationship between the total delay and ω is given in the following proposition.

Proposition 3. *There exists $\tilde{\omega}_1, \tilde{\omega}_2 \in (\frac{1}{2}, \frac{5}{6})$ such that (1) For $\omega \in [\frac{1}{2}, \tilde{\omega}_1) \cup [\tilde{\omega}_2, \frac{5}{6})$, the expected (total) delay is decreasing with respect to market concentration, as measured by the Herfindahl-Hirschman index. (2) For $\omega \in (\tilde{\omega}_1, \tilde{\omega}_2)$, the expected (total) delay is increasing with respect to market concentration.*

Principal Supplier. The Principal Supplier rule prescribes that countries should *focus* on negotiating with their major suppliers. In particular, I interpret the rule as one to determine which (foreign) country makes its offer first. Thus, in our bargaining context, the principal supplier rule simply means the principal supplier propose first. The asymmetry in my setting allows me to explore the role of this rule in the current bargaining context.

Proposition 4. *Under the given parametric setting, letting a Principal Supplier propose the initial offer makes the total (expected) delay shorter.*

This proposition provides a mechanism, different from Ludema and Mayda (2009, 2013), which lends additional support to the Principal Supplier rule embedded in the GATT/WTO institution.

Numerical Examples. As an illustrative example, I plot the numerical solutions for $h = 2$ and $h = 3$ respectively in the Figure 5.2, where the horizontal axis denotes the endowment of $\star 1$, and vertical axis is the amount of delay. In particular, the top 3 rows of panels represent the three components of the expected delay, plotted in the bottom row of panels. Moreover, I plot the outcome as red curves when $\star 1$ is the first proposer, and the outcome as blue curves when $\star 2$ is the first proposer.

The role of the principal supplier rule can be seen by viewing the bottom panel of the figure above: for $\omega > \frac{1}{2}$, that is $\star 1$ is the principal supplier, letting $\star 1$ proposes first (red curve) makes the expected shorter than letting $\star 2$ (smaller supplier) proposes first (blue curve).

While it is hard to directly measure externality in typical bargaining models, there is a popular candidate for it in trade settings. In general, the more concentrated the export market is, the smaller the scope for free riding (externality) is: in the extreme of only one supplying country, no free riding is present.[44] As can be seen from the numerical example presented in the bottom panels of the figure, delay is in general downward sloping with a kink, that is, the higher the market concentration is, the shorter the expected delay is.

Testable Predictions. While the model predicts many testable predictions (Proposition 1, 2 and 3), the first and preliminary one that I will take to the data is that given the principal supplier rule, the delay (from the initial offer to the final offer) will be

[44] Admittedly, I ignore the adjustment of countries' export pattern, as a lower tariff on a given good may induce a previously non-exporting country to become an exporter of that good.

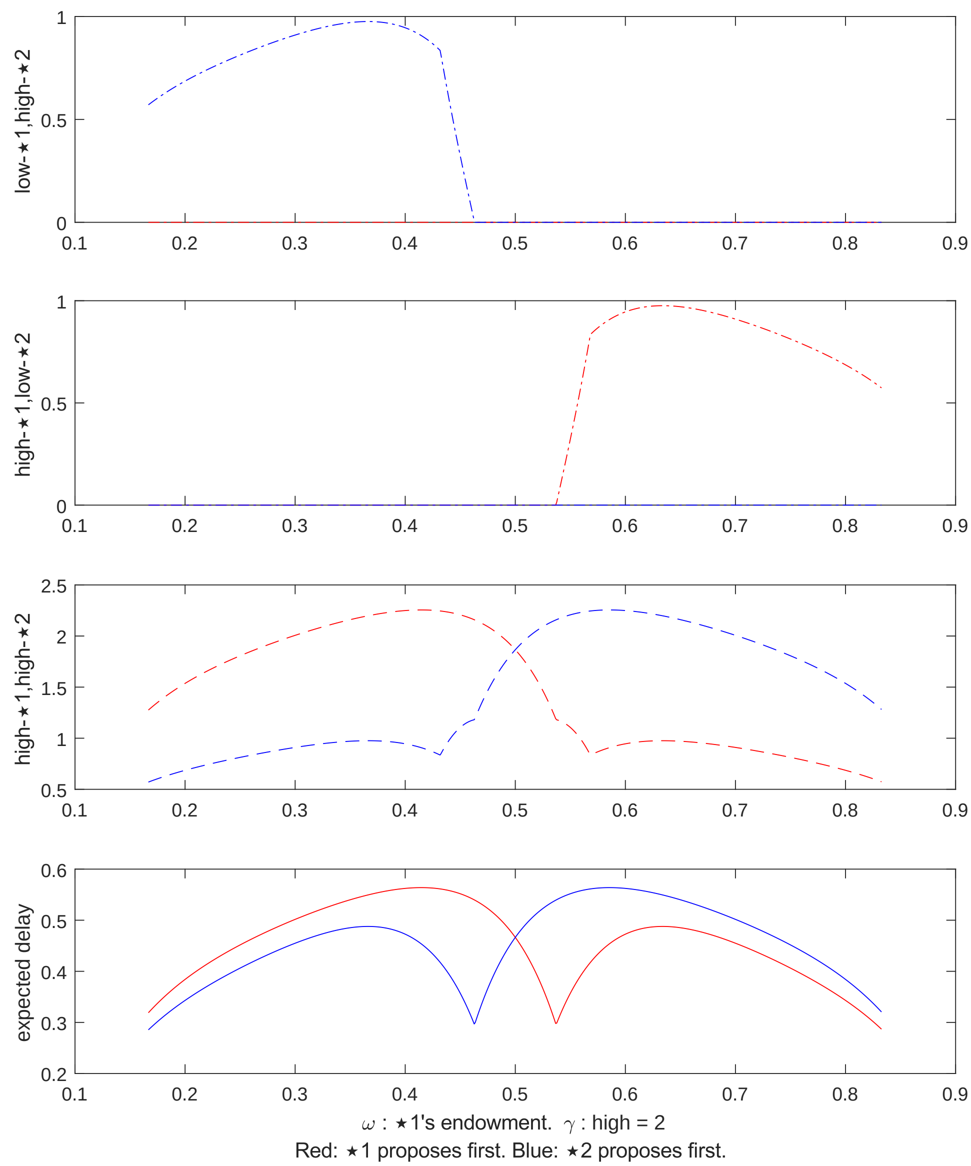

FIGURE 5.2.

maximized at $\omega = \frac{1}{2}$ (Proposition 1). Thus I expect that the delay will be decreasing with respect to the export concentration, that is, the more concentrated the export suppliers are, the shorter bargaining delay will be.

6. Empirical Results

6.1. **Data.**

Bargaining Data. I use the Torquay round data that is compiled and described by Bagwell, Staiger and Yurukoglu (2017). In particular, this dataset provides me detailed records of both the contents and the timing of offers, counteroffers and final agreements/failures between those negotiating countries. This round of tariff conference took place in Torquay, United Kingdom, on September 28, 1950, and lasted until the end of March 1951, where countries mainly focused on exchange of tariff concessions. Thus, in a sense, I do not need to deal with negotiations involving non-tariff measures that are typical in later rounds, for example, intellectual property rights. As a general bargaining procedure, countries bargain with each other in a *bilateral* manner, and they may do so *concurrently* with many partners. Prior to the arrival at Torquay, countries are advised to submit lists of products to other members on which they would like a partner to make concessions, that is, it is the list of concessions that they *request* for their exports into a partner's market. Once the round starts, countries begin exchanging their offers at whatever time that best suits them, and these intermediate offers are not observed by other third-party members, until they reach a final agreement.

Trade Data. The trade data in my analysis is the import data of United States in 1948, compiled and described by Bagwell, Staiger and Yurukoglu (2015).

Other Data. Another source of data for my covariates such as GDP, Common language, Border, Distance, comes from CEPII. And the market power measure comes from Broda Limao and Weinstein (2008).

6.2. **Empirical Strategy.** My model predicts that the delay from the first exporter offer to the final agreement is decreasing in the HHI calculated over the two exporters (i.e., ignoring the exports of non-bargainers). In the limit, as the HHI goes to 1, my model predicts delay from the first exporter offer to final agreement goes to zero. To take my model prediction to the data, in the sample of US data, I identify the US with the home country and I identify the foreign exporters bargaining with the home country over tariff i as the foreign exporters that the US makes an offer to on its tariff i (i.e., the foreign exporters that the bilateral bargaining records show the US offered to cut its tariff on a given product). And, in the sample of the full Torquay round data, I identify the importer as the home country for a given product-importer pair.

In total, there are 292 (undirected) dyads negotiating over thousands of products in the full Torquay sample.[45] The empirical specifications can be examined at the (directed or undirected) dyad level, although the relatively small number of observations in the dyad level limits the scope of the results. Alternatively, I can focus on the HS product level in two forms: (a) how long it takes to arrive at an agreement on the given HS product, regardless of the importing country (thus HS as the unit of observation). (b) how long it takes to arrive at an agreement on the given HS product with a particular importing country (thus HS-importer as the unit of observation). This second approach, in particular (b), not only greatly expands the number of observations, but also more naturally accords with the single-importer-multiple-exporter framework in the theory. However, there are some issues with these approaches which will be discussed in what follows.

A first question arises whether the procedure in the model is relevant to the actual negotiation, which affects how I can reasonably map the variables in the model to the data. In this regard, I argue the data can be interpreted in a way that is compatible with the theory. Specifically, I interpret the home country as the importing country in the data, and the foreign countries as the exporting countries. For a given product imported by the home country, when the home country makes an offer, I interpret this as corresponding to the moment that a foreign country confirms it is ready to

[45]Unless specifically mentioned, the numbers what follows are based on the full Torquay sample.

receive the home country's offer – and simultaneously makes its own offer to the home country.[46] In particular, I treat the home country's offer as if it is perfectly predicted by others, following the similar assumption made in the theoretical model where the home country's desire is known to everyone.[47] And, given the offer, the home country behaves passively (that is, it does not behave strategically in terms of delaying), and *officially* submits its offer to any other country whenever the latter *bids* for the offer with its own concession, guided by (multilateral) reciprocity. Consequently, given this interpretation, the exporters behave strategically by delaying their "bid", so as to signal their bargaining strength among themselves.

A second potential issue with linking the model to the data is the number of players, or countries, because in the data there are 36 negotiating parties,[48] nonetheless my model, if taken literally, has two foreign exporters with whom the home country bargains over the tariff on a given product. I argue, according to a strict interpretation of reciprocity in the negotiation (and together with MFN), there may exist other exporters of this product into the home market, but if they are not bargaining (and therefore do not alter their own trade taxes as a result of this bargain) then they will not be impacted by the bargaining between the home country and the two foreign countries – this reflects a result from Bagwell and Staiger (2005, forthcoming) that reciprocity plus MFN implies no third party impacts of a bilateral bargain. One feature of the data I use is that the majority (97%), as documented in Bagwell Staiger and Yurukoglu (2017), of product-level offers are made to no more than two countries. Given this observation, although introducing a more general setting with multiple countries might alleviate the discrepancy, it is beyond the scope of this paper and

[46] Given that the initial (final) offers of two countries are usually exchanged simultaneously or within a few days, this assumption serves as a good approximation in interpreting the actual data.

[47] Recall from footnote 22 that this assumption is not as restrictive as it seems, since there isn't much intensive margin adjustment in the offers, as documented in Bagwell, Staiger and Yurukoglu (2017).

[48] Some countries/regions were negotiating as a single party such as Benelux.

my focus on a two-exporter setting does not result in loss of generality in terms of the pattern observed in the data. In particular, in the data, a home country bargains with one foreign exporter on 79.2% of its tariffs; it bargains with two exporters on 17.5% of its tariffs; it bargains with three or more exporters on 3.3% of its tariffs. In particular, US offered tariff cuts to one explicit country on 52% of its trade volume that was under the Torquay negotiation, to two explicit countries on 38% of the volume and to three or more countries on 10% of the volume.

While the theory can be taken to the data in a way that makes the bargaining procedure sensible, delay in the actual negotiation can takes various forms, as there are several stages in the bargaining process: Request (R), Modification of Request (RM), Offer (O), Modification of Offer (OM), Final Agreement (A) and Modification of Final Agreement (AM), together with Termination/Withdrawn (W). Specifically, ignoring the modifications, there are three measures of delay that could serve as a potential focus of the empirical analysis.[49] The three measures of delay are the following: (a) the time elapsed from September 28, 1950, the official starting date of the Torquay round, to final offer, that is, the conclusion of an agreement i, which will be denoted as T_i^A; (b) the time elapsed from September 28, 1950, the official starting date of the Torquay round, to initial offer, which will be denoted as T_i^O; (c) the time elapsed from initial offer to initial final agreement, which will be denoted as T_i^{O-A}.

My main focus will be on (c) for the following reasons: (1) although the Torquay round was officially scheduled to start on September 28, 1950, it may not be a good measure of when the actual negotiation starts for each dyad: it may well be that some countries are not fully ready, exogenously; and (2) the theory only predicts a (simple) monotone relationship for (c).[50] While I focus on the initial offer and initial final offer,

[49]The missing data problem in the request stage is significant, as many countries do not report the date of request - at least not in a way that I can observe.

[50]While I do see sign changing in different specifications related to T_i^A and T_i^O in my analysis, a more extensive examination of the non-monotonicity in (a) and (b) is left for future work.

missing the modifications of offers in between, the average number of back and forth of offers is close to two, as documented in Bagwell, Staiger and Yurukoglu (2017). In particular, a majority of bilateral bargainings only involves 2 rounds of offers, that is, countries started the initial exchange of offers (O), then finalized the offers (A) at a later time - the elapsed time can be interpreted as observations of t_{hh}, consisting of t_{hl} (elapsed time from beginning of round to O) and $t_{hh} - t_{hl}$ (elapsed time from O to A), in my theoretical model. Then the second most frequent number of rounds of offers is 1, which consists of two scenarios: (i) countries start their exchange of offers (O), then nothing occurred. (ii) countries reach an agreement (A) without initial exchange of offers. For case (i), I can treat them as censored observations where the delay is censored at the end of round.[51] For case (ii), the elapsed time from the beginning to (A) can be mapped to t_{hl} in my theoretical model. In total, these forms account for around 75% of the bilateral bargainings in the Torquay round.

As the variable of key interest, the bargaining externality can be proxied indirectly using the number of exporters involved in each bargain with an importer. However, this measure could be endogenous to the bargaining setting (as in Ludema and Mayda, 2013), and beyond this the use of such a proxy might lead to a model that is hard to distinguish from a mechanical one, say, a random proposing model where the more players there are, the earlier the initial (earliest) offer will be made and the later the initial (earliest) final agreement will occur, which mechanically maps larger externality (more players) to longer delay. While the latter should not concern the current setting, as there is only one importer making offers, another (better) approach is to measure the externality directly. One such measure would be the product concentration in a given country/importer, where the common HHI can be applied. The potential problem with this approach is how to measure the concentration. Specifically, I have

[51]Note that it is evident that bargaining failure is part of the dataset, which, admittedly, my model cannot explain *per se*. As a result, I have to exclude the bargaining pairs that failed in my sample, although this portion of data is relatively small.

to determine whether all exporters, or only those that participate in the negotiation, should be included in the measure. In this regard, my theoretical framework provides a remedy. In particular, as I impose reciprocity (going-down), together with MFN, according to the model there should be no free rider issue related to countries that do not come to the bargaining table (i.e. those non-participants would not benefit from any agreement). Thus, guided by this observation, I should focus on the HHI among those actual negotiators for a given HS.

At first glance, my delay framework seems similar to the analysis of strike duration in labor economics, where hazard rate/survival model is usually applied.[52] However, this modeling strategy is not necessarily applicable in my setting. In hazard modeling, the hazard rate is typically assumed to be constant, which can be rationalized by mixed strategies - thus, in a sense, this type of model embodies the feature of war of attrition, which my theoretical model does not have as I focus on pure strategy.[53] In particular, length of delay in my theoretical framework is itself a strategic device that can be directly linked to other variables, as opposed to a setting where delay is a by-product of hazard rate. As a result, I will focus on other specifications such as simple OLS and Tobit.[54]

My initial specification will be taking Proposition 1 to the data, that is, to examine whether delay is related to externality. As shown in the empirical evidence that follows, the signs are in general consistent with this proposition in all tables, that is, the larger the externality is, the longer delay (from initial offer to final offer) will be. Also, the theory (Proposition 2 and Proposition 3) does not predict a monotone pattern with respect to the relationship between bargaining delay (from beginning

[52]See Kennan (1985), Tracy (1987) and Gu and Kuhn (1998), among others, for this line of modeling. Moser and Rose (2012) employ survival analysis to examine delay in trade negotiations across different regional trade agreements.

[53]And indeed, the hazard model itself is rejected by the data.

[54]Neary (2004) fits an OLS relationship between the duration of negotiation and the number of participating countries, across different GATT rounds, and uses his estimates to form a prediction of the duration of Doha round.

Variable	N	Mean	SD	Min	Max
Days from 9/28/1950 to Initial Offer	324	56.06	54.43	-20	184
Days from Initial Offer to Final Offer	260	80.86	59.82	0	199
Days from 9/28/1950 to Final Offer	260	134.4	50.33	11	184

TABLE 1. Descriptive Statistics: Dyad Level

Variable	N	Mean	SD	Min	Max
Days from 9/28/1950 to Initial Offer	1,760	35.68	41.21	4	184
Days from Initial Offer to Final Offer	1,267	142.6	43.32	0	179
Days from 9/28/1950 to Final Offer	1,267	172.3	15.28	77	184
US Import HHI	1,760	0.761	0.372	0	1

TABLE 2. Descriptive Statistics: HS Level (US Import)

to initial offer, or from beginning to final agreement) and externality, and I do see some switching of coefficient signs in different specifications. In particular, my basic regression equation is:

$$T_i = \alpha + \beta EXT_i + \boldsymbol{X}_i' \boldsymbol{\gamma} + \epsilon_i$$

where T_i is my measure of delay - in particular, the delay from an initial offer to final agreement, EXT_i is the measure of externality and $\boldsymbol{X}_i$ is the vector of control variables. Based on my theoretical results in Proposition 1, one would expect $\beta > 0$. I summarize the data in Table 1 and Table 2. And I present the results in two sets for different samples in what follows.

6.3. **Results: Sample of US Bilateral Bargaining.** Trade data definitely served as the basis for the tariff negotiations in Torquay and was arguably the most important information that countries possessed.[55] Since only the US import data, for the required time period and at the required level of disaggregation, is currently available, I focus on the sub-sample where the US was a seller (importer) in the negotiations.[56]

[55]The GATT advised countries to exchange their lists of products under consideration, together with trade statistics.

[56]I will also analyze the bargaining data for all Torquay countries later with a measure of externality that does not require import data.

As previously discussed, the key variables that I have to measure, or define, are T_i, the length of negotiating delay, and EXT_i, the degree of bargaining externality. Thus I proceed according to various definitions of EXT_i, and present the results for different T_i, given the definition of EXT_i. Moreover, the unit of observation i, could be either in the product level, or it can be based on the dyad level where a bundle of products are under negotiation. Thus I divide my results into two groups based on how the unit of observation is defined.

Unit of Observation: 6-digit Harmonized System product level (HS product)

I first make use of the trade data and relate the externality to the export market concentration, which is a direct result of theory. In particular, I relate the externality to the HHI of a given HS among negotiating exporters in the US market, characterizing how concentrated the export market is for a given HS product.[57] Column (1) and (2) of Table 3 present the results using this direct measure of externality. While a complete set of Torquay data is desirable, focusing on US can serve as my initial examination of the externality issue in a direct way. In particular, they are statistically significant, and seem to be consistent with Proposition 1: the smaller the exporter concentration is (thus larger externality), the longer the delay is.

Another way to measure externality is based on the number of partners that are negotiating on US's tariff concession, as my theory, although trivially, predicts there should be no delay when there is only one partner on a given product, while delay starts to emerge as more partners are involved. Note that this measure does not require knowledge of the import trade volume, allowing for a full Torquay sample analysis whose results will be reported in the next subsection. For comparison purpose, I also report the results for the US sample here. In particular, I relate the

[57]Concretely, denoting $\mathbf{C}^{US} \equiv \cup \mathbf{C}_i^{US}$, where $\mathbf{C}_i^{US} \equiv$\{countries that had received offers on product i from US\}, I define $EXT_i \equiv 1 - HHI_i$ with $HHI_i \equiv \sum_{c \in \mathbf{C}_i^{US}} s_{ci}^2$, where s_{ci} is country c's existing share in the export market of product i to US. In the data, there are offers made to countries on goods that none of them exports. I simply exclude them as exogenous data error.

	HHI		Single *vs* Multiple	
VARIABLES	(1)	(2)	(3)	(4)
	OLS	Tobit	OLS	Tobit
EXT	26.19***	15.69**	21.78***	16.24***
	(7.985)	(7.994)	(2.471)	(2.501)
Constant	136.3***	140.6***	130.2***	136.0***
	(1.814)	(1.822)	(1.878)	(1.893)
Observations	1,017	1,061	1,267	1,328
(Pseudo) R-squared	0.005	0.000	0.032	0.002
Standard errors (robust) in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

TABLE 3. Larger externality leads to longer delay.

externality to whether a good is negotiated with multiple countries, that is, EXT_i is defined as a dummy variable which takes 1 if there are more than one country that had received offers on product i from US, and 0 otherwise. The results under this alternative measure are presented in column (3) and (4) of Table 3. As can be seen, the externality coefficients are statistically significant, and their signs are consistent with what the theory would predict, that is, larger externality leads to longer delay.

However, a potential issue with these specifications is the "bundling" of products in the negotiation, which may cause correlation between observations. For example, if two products are negotiated with the same exporter(s), then the delay on the two products becomes identical. To accommodate this issue, following Cameron and Miller (2015) and based on the definition of delay – from initial offer to final agreement, I control for the within-cluster correlation by adding in the cluster-specific fixed effects and also clustering the standard errors where a cluster is defined by the identity of the exporting country receiving the initial offer and the identity of the exporting country receiving the final agreement.[58] As presented in Table 9, the results are similar to those in Table 3.

[58]As pointed out by Cameron and Miller (2015), cluster-specific fixed effects alone may not fully control for the within-cluster correlation. For comparison purpose, I also report the results with cluster-specific fixed effects, but without clustering the standard errors. However, the results with

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	13.32*	13.32**	6.28	6.28	11.96***	11.96***	6.86**	6.86***
	(7.500)	(5.882)	(7.397)	(5.538)	(2.774)	(2.701)	(2.762)	(2.426)
Constant	147.7***	147.7***	152.2***	152.2***	144.8***	144.8***	149.2***	149.2***
	(1.814)	(0.998)	(3.585)	(0.954)	(1.608)	(1.182)	(2.624)	(1.013)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.501	0.501	0.063	0.063	0.463	0.463	0.053	0.053
FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

TABLE 4. Larger externality leads to longer delay.

Unit of Observation: Dyad

As an alternative to the product level negotiation, I can measure the delay at the dyad level, that is, how long it takes a dyad to conclude their negotiation, if any, since the timing of negotiation of a dyad is formally based on the list of products that are under negotiation.

In principle, the importance, and difficulty, of a bargaining or an agreement depends on what is at stake. Intuitively, it would be ambiguous to compare the difficulty facing a high-concentration good with large trade volume to the difficulty facing a low-concentration good with little trade volume. In particular, when countries exchange their list of offers, goods on the same list might not be of equal importance in terms of their respective trade volume. Thus, I relate the externality for a dyad to the *average* HHI across the products under their negotiation, weighted by their import share. In particular, denoting the negotiation list of dyad i as $\mathbf{J_i}$ and product j's import share as w_j, I define $EXT_i \equiv 1 - \sum_{j \in \mathbf{J_i}} w_j HHI_j$. Table 5 presents the results. And I do see patterns in this table that suggest a correct sign of the externality coefficient, although several of them are not very significant. In particular, I find in both regression (1) and (2) a positive association between externality and the bargaining delay.

clustering the standard errors, but without cluster-specific fixed effects are not reported due to the concern that clustering the standard errors alone may not fully control for the correlation either.

	HHI (average)	
VARIABLES	(1)	(2)
	OLS	Tobit
EXT	345.5*	398.9**
	(162.9)	(181.0)
Constant	-199.4	-224.7
	(155.8)	(172.7)
Observations	15	20
(Pseudo) R-squared	0.249	0.018
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

TABLE 5. Externality measured by concentration of the export market: averaging.

Discussion: home short

The key part of the theory is built on the cases of home short (but not too short), namely being close to (weakly) home short. And, as discussed in the Appendix, in the home long case, there should be no delay. To further examine this in the data, I proxy the home short/long position by measuring the gap between a request and an initial offer on a given HS6 product. To accommodate the fact that there is no formal request stage in the theory, I adopt the following simplifying (plausible) assumption: the request stage occurs before any (unexpected) political shocks take place.[59] As a result, interpreted through the lens of my model, it is as if countries make their requests when both of them are of low political pressure, that is, the status quo.

[59]More formally, the setup can be outlined as follows. There are two *unexpected* shocks: one (the first shock) is the shock that allows the countries to start a negotiation, and the other (the second shock) is the political shock which determines each country's new domestic political pressure. Following the standard literature on political pressure, I interpret the lobbying activity as the source of political pressure. In particular, countries exchange their requests after the first shock, and lobbying (unexpectedly) takes place after they submit their lists of requests (perhaps because only after requests have been submitted do industries know what products will be the focus of negotiations). As a result, information in the request stage is complete – each country's political pressure is common knowledge, yielding a standard alternating bargaining game as perceived by the participating countries in the request stage. Then, before any agreement is formally signed on, lobbying unexpectedly occurs, resulting in re-evaluation of their positions by each country. Since the consequent political pressure is privately observed, the subsequent offer stage becomes an alternating bargaining game with incomplete information where strategic delay becomes possible.

In particular, the gap is defined as the ratio of (initially) offered tariff rate to the requested rate, which I denote as r. Then I can discuss the following cases: (a) when $r > 1$, that is, the offered rate is larger than the requested rate, it implies that the home country is short, which wants less concession. (b) when $r < 1$, that is, the offered rate is smaller than the requested rate, it implies that the home country is long, which wants more concession. (c) when $r = 1$, that is, the offered rate is equal to the requested rate, it implies that the home country is equal, which wants exactly what the foreign countries desire. However, under the assumption that the requested rate is made when both of the two foreign countries are of low type, the home is equal when both of the two foreign countries are of low type. Thus among the cases discussed in the Appendix, the home equal case corresponds to the weakly home short case, instead of the weakly home long case. And we simply categorize $r \geq 1$ as the home short case, and $r < 1$ as the home long case.[60] Based on this definition, I divide the sample into two groups, between which I examine the difference in delays, and the difference in the regression outcomes, as a further validity check to the theory. Specifically, I focus on two aspects that the theory predicts: (1) there should be no delay in the home long case which I interpret as implying that the delay in the home long group should be shorter than the delay in the home short group. (2) it should be observed that there exists stronger monotonically increasing relationship between delay and externality in the home short case.[61]

(1) Delay: home short *vs* home long

[60] Admittedly, a significant portion of observations is lost in the process, due to the fact (1) that the tariff reduction requested by Benelux is merely "intent" without specified rate – thus r is not well-defined, (2) that some products are not in the request list, that is, they are added in the offer stage.

[61] Although non-monotonicity, in the home short case when the home is too short, is itself an interesting pattern, it is ambiguous to define how short is "too short": after trying different cutoffs for home shortness, I find that the monotonically increasing relationship holds in these sub-samples, that is, non-monotonicity is not found in the sample. One possible explanation is that the maximum shortness in the sample is still below a threshold that makes home short enough to examine the non-monotonicity.

	Home Short		Home Long	
VARIABLES	(1)	(2)	(3)	(4)
	OLS	Tobit	OLS	Tobit
EXT	37.32***	28.23***	-14.42	-18.36
	(8.050)	(8.009)	(25.565)	(25.497)
Constant	139.2***	142.9***	136.1***	137.9***
	(2.914)	(2.916)	(5.476)	(5.475)
Observations	333	346	121	123
(Pseudo) R-squared	0.017	0.001	0.002	0.000
Standard errors (robust) in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Table 6. Home Short *vs* Home Long

As a first pass, I simply compare the mean delay of the two groups: the mean delay of the home short group is 142.4 days, while the mean delay of the home long group is 134.4 days. While the difference of approximately one week seems economically insignificant, a simple t test shows that the former is larger than the latter at the 5% significance level.[62]

(2) Monotonically increasing relationship

To further investigate this implication of my model, I conduct the same regressions as in the regression (1)-(2) of Table 3 for the two groups – home short, and home long – separately. The results are presented in Table 6 and Table 7. As can be seen, the monotonically increasing relationship does hold in the home short group. However, such a relationship is not present in the home long group, as indicated by the insignificant coefficients.

6.4. **Results: Full Sample of Torquay Negotiation.** In this subsection, I focus on a measure of the externality that does not require knowledge of the import trade volume (namely, the indicator variable defined above that takes the value of 1 if there

[62]If taken more literally, the theory predicts that there should be no delay for the home long group. To directly check this, I compare the ratios of 0-delay between the two groups and find that the ratio in the home long group is slightly larger.

	Home Short				Home Long			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	20.05***	20.05***	10.37	10.37**	-20.18	-20.18	-20.18	-20.18
	(6.669)	(5.823)	(6.517)	(4.527)	(31.748)	(25.868)	(29.411)	(23.965)
Constant	146.2***	146.2***	148.0***	148.0***	155.2***	155.2***	155.2***	155.2***
	(1.934)	(1.112)	(1.438)	(0.864)	(8.346)	(6.685)	(7.732)	(6.194)
Observations	333	333	346	346	121	121	123	123
(Pseudo) R-squared	0.475	0.475	0.059	0.059	0.518	0.518	0.061	0.061
FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

Table 7. Home Short *vs* Home Long

was more than 1 exporting country involved in the negotiation on a given import tariff, and zero otherwise), and hence allows me to use the complete Torquay bargaining data, which extends the US sample that I used in Section 6.3 to all Torquay bilateral bargaining. In this larger sample, every country is considered as the home country when considering its own tariff concession, and also treated as a foreign country when considering the tariff concession offered by any of its partners.

Unit of Observation: HS-importer

A product is defined as a "HS-importer" pair, which is the unit of observation I hereby focus on (e.g. I distinguish between a product imported by US and the same product imported by Canada). And the delay is measured as how long it takes to reach an agreement, if any, on a given product. Since the period import data for all countries that participated in the Torquay Round is not currently available, here the externality is, again, measured by considering whether a given HS-importer is negotiated with a single buyer or multiple buyers. Table 8 presents the preliminary results. In particular, regression (1) and (2) focus on the full sample of Torquay, and the French offers are excluded from regression (3) and (4). While the coefficient from OLS (1) is highly significant and its sign is consistent with the theoretical prediction – that is, a larger externality leads to a longer bargaining delay – Tobit (2) has an

	Single *vs* Multiple			
	Full Sample		w/o French imports	
VARIABLES	(1)	(2)	(3)	(4)
	OLS	Tobit	OLS	Tobit
EXT	38.42***	-1.890	40.03***	3.585*
	(1.514)	(1.851)	(1.646)	(2.015)
Constant	63.5***	118.6***	66.7***	119.7***
	(0.777)	(1.119)	(0.827)	(1.165)
Observations	11,123	14,721	9,685	12,830
(Pseudo) R-squared	0.050	0.000	0.050	0.000
Standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

TABLE 8. Externality measured by whether the product is negotiated with a single buyer or multiple buyers.

insignificant parameter with the opposite sign. Upon further examination, I find that the Tobit results from regression (2) are driven largely by the offers made by France. Indeed, as argued in Bagwell, Staiger and Yurukoglu (2017), France is special in terms of the strategies it used during the Torquay round, which also significantly affects some of their empirical results. Thus, I present the results after excluding French offers in regression (3) and (4). The coefficients of interest in both specifications are significant and consistent with the sign the theory predicts. Moreover, comparing the results to Section 6.3, it also suggests that there might be some heterogeneity across different countries (importers). Also, the results after after controlling for the within-cluster correlation are presented in Table 9. While the OLS specifications exhibit similar results, significance is lost in Tobit (4) and (8).[63]

Unit of Observation: Dyad

[63]While the loss of significance lends no support to the theory in the paper, it is worth noting that the results from a tobit model with fixed effects should be taken with caution, due to the so-called "incidental parameters" problem. See Greene (2004) for the discussion of the issues in this type of econometric modeling.

	Single *vs* Multiple							
	Full Sample				w/o French imports			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	20.34***	20.34***	-7.42***	-7.42	18.95***	18.95***	-8.48***	-8.48
	(1.808)	(3.220)	(2.485)	(5.403)	(1.858)	(3.105)	(2.734)	(5.969)
Constant	30.3*	30.3***	106.1***	106.1***	30.3*	30.3*	106.2***	106.2***
	(17.124)	(0.000)	(34.286)	(0.996)	(17.149)	(0.000)	(34.353)	(1.051)
Observations	11,123	11,123	14,721	14,721	9,685	9,685	12,830	12,830
(Pseudo) R-squared	0.511	0.511	0.060	0.060	0.536	0.536	0.056	0.056
FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

Table 9. Externality measured by whether the product is negotiated with a single buyer or multiple buyers.

Although I currently cannot rely on the import data for the full sample, an analysis that might be interesting on its own is how delay is related to the number of negotiating partners, that has been examined in other literature such as Moser and Rose (2012) and Neary (2004).[64] In particular, for a dyad ij, with i being the seller and j being the buyer, I examine the relationship between delay and (a) the number of buyers facing seller i ; (b) the number of sellers facing buyer j. Table 10 presents the result. First, by focusing on variations of delay *within* a round, I see similar evidence to those examined in Moser and Rose (2012) and Neary (2004), which instead are based on variations of delay across agreements/rounds. Namely, larger number of participants is associated with longer bargaining delay. Second, this set of evidence is, in a sense, consistent with the idea regarding externality and bargaining delay. In particular, when there are more countries *involved* in the negotiation, the bargaining delay will be longer.

6.5. **Robustness and Sensitivity.** As an initial check, I introduce the country-fixed effects into the specifications introduced in the previous section, to account for

[64]Admittedly, my two-exporter framework has little to say about the relationship between delay and the number of exporters.

	Number of Countries	
VARIABLES	(1)	(2)
	OLS	Tobit
#Buyers	3.132***	2.148**
	(0.768)	(1.065)
#Sellers	2.357***	0.902
	(0.661)	(0.979)
Constant	1.893	69.93**
	(20.67)	(29.91)
Observations	260	324
(Pseudo) R-squared	0.090	0.002
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 10. Externality measured by the number of "players".

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	39.65***	39.65**	39.09***	39.09**	21.18***	21.18*	20.51***	20.51*
	(11.93)	(17.141)	(12.100)	(17.602)	(6.817)	(11.020)	(6.999)	(10.986)
Constant	122.2***	122.2***	131.0***	131.0***	126.5***	126.5***	136.0***	136.0***
	(3.841)	(6.469)	(3.864)	(6.681)	(5.416)	(9.922)	(5.584)	(9.362)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.352	0.352	0.038	0.038	0.322	0.322	0.033	0.033
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

Table 11. Larger externality leads to longer delay.

unobservable country difference that might affect the bargaining delay. In particular, Table 11 adds the country-fixed effects into Table 3. The results are qualitatively similar.

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	35.29*	35.29*	30.04**	30.04**	24.54***	24.54**	18.40***	18.40**
	(19.090)	(20.293)	(13.940)	(14.920)	(8.507)	(11.146)	(6.360)	(7.521)
Constant	160.4***	160.4***	165.8***	165.8***	161.3***	161.3***	166.3***	166.3***
	(8.686)	(8.659)	(6.457)	(5.581)	(8.917)	(9.652)	(6.623)	(5.812)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.892	0.892	0.210	0.210	0.866	0.866	0.189	0.189
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
HS4 FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

TABLE 12. Larger externality leads to longer delay.

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	55.81***	18.78*	46.67***	12.04	24.20***	7.790	17.84***	2.43
	(16.065)	(9.903)	(11.653)	(7.443)	(5.051)	(5.271)	(3.830)	(2.211)
Constant	176.0***	166.4***	176.0***	170.3***	176***	161.5***	176	168.5
	(0.000)	(10.258)	(0.000)	(7.168)	(0.000)	(10.988)	(.)	(5.855)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.870	0.929	0.196	0.245	0.839	0.909	0.174	0.222
HS4 FE	Y	Y	Y	Y	Y	Y	Y	Y
Cluster FE		Y		Y		Y		Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

TABLE 13. Larger externality leads to longer delay.

Also, another potential concern is the unobservable levels of political pressure,[65] which might cause confounding effects to my specifications. I control this by respectively using the HS4 and HS2 fixed effects with or without the country-fixed effects. The results presented in Table 12 and Table 15 are also similar.

[65]In principle, I could make use of political contribution to proxy for the pressure, which Goldberg and Maggi (1999) use to identify an *organized* industry. Not without its own issue as discussed in Goldberg and Maggi (1999) and Gawande and Krishna (2004), this approach would require a much larger dataset of political contribution from multiple countries. Moreover, this data has to be near the time of Torquay conference. Gawande, Krishna and Olarreaga (2015) estimate the political weights in a cross-country manner based on data over the 1988-2000 period, which seems too far from the 1950-1951.

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	51.74***	51.74***	47.76***	47.76***	23.16***	23.16*	22.26***	22.26**
	(11.580)	(16.030)	(11.320)	(15.320)	(6.556)	(11.857)	(6.560)	(10.764)
Constant	144.0***	144.0***	160.1***	160.1***	145.1***	145.1***	165.1***	165.1***
	(5.404)	(5.832)	(9.463)	(14.045)	(6.798)	(7.808)	(10.780)	(12.350)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.521	0.521	0.066	0.066	0.468	0.468	0.054	0.054
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
HS2 FE	Y	Y	Y	Y	Y	Y	Y	Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

Table 14. Larger externality leads to longer delay.

	HHI				Single *vs* Multiple			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
EXT	36.62***	7.71	24.20***	-0.87	19.28***	8.54**	12.41***	2.65
	(8.989)	(10.426)	(8.604)	(8.728)	(2.939)	(3.539)	(2.895)	(3.299)
Constant	155.1***	158.3***	171.1***	165.5***	153.2***	152.3***	170.7***	164.9***
	(5.324)	(3.985)	(11.131)	(14.443)	(6.211)	(4.756)	(11.787)	(14.624)
Observations	1,017	1,017	1,061	1,061	1,267	1,267	1,328	1,328
(Pseudo) R-squared	0.357	0.660	0.040	0.096	0.305	0.591	0.031	0.076
HS2 FE	Y	Y	Y	Y	Y	Y	Y	Y
Cluster FE		Y		Y		Y		Y
CE		Y		Y		Y		Y
Standard errors (robust) in parentheses								
*** p<0.01, ** p<0.05, * p<0.1								

Table 15. Larger externality leads to longer delay.

7. Conclusion

This paper studies bargaining delay in the environment of multilateral trade negotiations. A simple bargaining structure is incorporated into an equilibrium trade framework (Section 3), which is used to establish the relationship between delay in reaching agreements and bargaining externality (Section 4 and Section 5). It is shown that the delay does vary systematically with the degree of export market concentration. And along this dimension, it is also argued that the Principal Supplier rule helps to reduce bargaining delay. Moreover, it is found that the empirical evidence based on the bargaining data from the GATT (Torquay round), due to its recent availability, lends some support to the theory (Section 6). One interesting implication of the theory relates to an implication of the formation of custom unions. In terms of bargaining delay, such integration helps to increase market concentration, thus reduce bargaining delay.

The analysis has made only an initial attempt at understanding the phenomenon of delay in reaching agreements plaguing most trade negotiations, which not only leads to potentially significant welfare loss, but also may hinder the momentum of trade liberalization, making the existing trade institution in danger. Many issues remain to be explored. First, most trade negotiations – especially within the context of the GATT/WTO, whether the simple bargaining structure imposed in the paper captures sufficiently well the key aspects of the actual bargaining process is not examined, as any (bilateral or not) bargaining is, in fact, part of a much broader *network* of players. Second, equally, if not more, important is the multiplicity of products/issues under consideration, where one would conjecture that bargaining on one good may be interdependent with another. Finally, the institutional rules such as MFN and Reciprocity are taken as given in the paper, an interesting and important question is how these rules, individually or jointly, affect the bargaining delay.

8. Appendix

The appendix contains the derivations and proofs omitted in the main text.

A. Parameter restriction

This subsection solves the trade equilibrium and discusses the imposed restrictions on the parameter space.

Since $Q^h = 0$, $Q^{\star 1} = \omega$, $Q^{\star 2} = 1 - \omega$. $D^h = 1 - \alpha p^h$, $D^{\star 1} = \beta - p^{\star 1}$, $D^{\star 2} = \beta - p^{\star 2}$. the only *ex ante* asymmetry comes from the different endowments in the foreign countries. For any given set of trade policies $(\tau^h, \tau^{\star 1}, \tau^{\star 2})$, the home country's import demand function is

$$M^h(p^h(\tau^h, p^w)) = 1 - \alpha p^h = 1 - \alpha(p^w + \tau^h)$$

and the (total) export supply is given by

$$\sum_{j=1,2} E^{\star j}(p^{\star j}(\tau^{\star j}, p^w)) = 1 - (2\beta - p^{\star 1} - p^{\star 2}) = 1 - 2\beta + 2p^w + \tau^{\star 1} + \tau^{\star 2}$$

Thus, market clearing implies $p^w = \frac{2\beta - \alpha\tau^h - \tau^{\star 1} - \tau^{\star 2}}{2+\alpha}$. [66]

In the initial period, countries impose their Nash tariffs. In particular, the Nash Equilibrium can be written as

$$\begin{aligned}
\tau_0^h =& \frac{1 - \alpha p_0^w}{2 + \alpha} \\
\tau_0^{\star 1} =& \frac{(1+\alpha)(\gamma_0^{\star 1} - 1)\omega + \beta - p_0^w - \omega}{2 + \alpha} \\
\tau_0^{\star 2} =& \frac{(1+\alpha)(\gamma_0^{\star 2} - 1)(1-\omega) + \beta - p_0^w - (1-\omega)}{2 + \alpha}
\end{aligned}$$

[66] In an endowment economy with specific export subsidies, presence of any political pressure will effectively make an export subsidy almost equivalent to a transfer to the (politically weighted) export sector. I will not obtain this corner solution in frameworks with more general supply system or those with ad valorem export subsidies. Also, in frameworks with import policies rather than export policies, for example, Grossman and Helpman (1995), the import tax makes the interior Nash equilibrium possible, regardless of whether an endowment economy is used. Thus I will impose the condition that all countries are national income maximizers, that is $\gamma_0^k = 1$, in the initial period, which will ensure an (interior) existing trade policies in Nash equilibrium or bliss point.

implying

$$p_0^w = \frac{2\beta + 1 + (2\beta - 1)\alpha + (1+\alpha)(1-\gamma_0^{\star 1})\omega + (1+\alpha)(1-\gamma_0^{\star 2})(1-\omega)}{2+4\alpha}$$

Since I assume they are initially all national income maximizers ($\gamma_0^{\star 1} = \gamma_0^{\star 2} = 1$), I must have

$$\begin{aligned}
p_0^w &= \frac{2\beta + 1 + (2\beta - 1)\alpha}{2+4\alpha} \\
\tau_0^h &= \frac{[2 - (2\beta - 1)\alpha](1+\alpha)}{(2+4\alpha)(2+\alpha)} \\
\tau_0^{\star 1} &= \frac{(2\beta+1)\alpha - 1 - (2+4\alpha)\omega}{(2+4\alpha)(2+\alpha)} \\
\tau_0^{\star 2} &= \frac{(2\beta+1)\alpha - 1 - (2+4\alpha)(1-\omega)}{(2+4\alpha)(2+\alpha)}
\end{aligned}$$

implying

$$\begin{aligned}
p_0^h &= \frac{(2\beta+1)(1+\alpha) + 1}{(1+2\alpha)(2+\alpha)} \\
p_0^{\star 1} &= \frac{(2\beta-1)\alpha^2 + 8\alpha\beta + 4\beta + 1 - (2+4\alpha)\omega}{(2+4\alpha)(2+\alpha)} \\
p_0^{\star 2} &= \frac{(2\beta-1)\alpha^2 + 8\alpha\beta + 4\beta + 1 - (2+4\alpha)(1-\omega)}{(2+4\alpha)(2+\alpha)}
\end{aligned}$$

Also, the equilibrium (initial) import/export volume is

$$\begin{aligned}
M_0^h &= 2\tau_0^h = \frac{[2 - (2\beta - 1)\alpha](1+\alpha)}{(1+2\alpha)(2+\alpha)} \\
E_0^{\star 1} &= -(1+\alpha)\tau_0^{\star 1} = -(1+\alpha)\frac{(2\beta+1)\alpha - 1 - (2+4\alpha)\omega}{(2+4\alpha)(2+\alpha)} \\
E_0^{\star 2} &= -(1+\alpha)\tau_0^{\star 2} = -(1+\alpha)\frac{(2\beta+1)\alpha - 1 - (2+4\alpha)(1-\omega)}{(2+4\alpha)(2+\alpha)}
\end{aligned}$$

implying $E^{\star 1} > E^{\star 2} \Leftrightarrow \omega > \frac{1}{2}$. In other words, the country with a larger endowment will be the *principal supplier*, referring to the foreign country that is the largest export supplier to the home country of the good under consideration. As will be seen in my bargaining outcome, this parametrization allows me to examine the effect of another institutional rule in GATT/WTO, in addition to MFN and reciprocity in negotiations, namely the Principal Supplier rule which leads countries to negotiate concessions with a principal supplier on a given product. In my bargaining model, I will capture this institutional feature by allowing the principal supplier to make the first offer, and I will compare the delay under this rule to the alternative in which the smaller supplier makes the first offer. Also, in order to preserve the trade pattern in the initial equilibrium, namely (1) positive export volume by $\star 1$ and $\star 2$, (2) positive domestic consumption in $\star 1$ and $\star 2$, I will restrict myself to $\omega \in \Omega(\alpha, \beta)$, where $\Omega(\alpha, \beta)$ is defined as:[67]

$$\underbrace{[\frac{(2\beta+1)\alpha-1}{2+4\alpha}, 1-\frac{(2\beta+1)\alpha-1}{2+4\alpha}]}_{(1)} \cap \underbrace{[(1+\alpha)\frac{1-(2\beta+1)\alpha}{2+4\alpha}, 1-(1+\alpha)\frac{1-(2\beta+1)\alpha}{2+4\alpha}]}_{(2)}$$

which reduces to

$$[\max\{\frac{(2\beta+1)\alpha-1}{2+4\alpha}, (1+\alpha)\frac{1-(2\beta+1)\alpha}{2+4\alpha}\}, 1-\max\{\frac{(2\beta+1)\alpha-1}{2+4\alpha}, (1+\alpha)\frac{1-(2\beta+1)\alpha}{2+4\alpha}\}]$$

To ensure $\Omega(\alpha, \beta) \neq \emptyset$, I must have $(2\beta-1)\alpha \leq 2$. In addition, to ensure a positive world price p_0^w, I also require $(2\beta-1)\alpha > -(2\beta+1)$. However, on the one hand, if $2\beta < 1$, these two conditions imply $\alpha > \frac{1+2\beta}{1-2\beta} > 1$, excluding the interesting case when $\alpha \leq 1$, which, as I will discuss later, is the case when home is *short*. On the other hand, if $2\beta > 1$, these two conditions imply $\alpha \leq \frac{2}{2\beta-1}$, which excludes the case $\alpha > 1$ for large β. To allow for both of the two possibilities, namely, $\alpha \leq 1$ and $\alpha > 1$, I restrict myself to $\beta \in [\frac{1}{2}, \frac{3}{2})$.[68]

The bliss points can now be characterized as $\bar{\tau}_{po}^{\star 1} = (\gamma^{\star 1}-1)\omega$, $\bar{\tau}_{po}^{\star 2} = (\gamma^{\star 2}-1)(1-\omega)$, $\tau_{po}^{h} = 0$, $p_{po}^{w} = \frac{2\beta+(1-\gamma^{\star 1})\omega+(1-\gamma^{\star 2})(1-\omega)}{2+\alpha}$. Given the restriction $p^{\star j} \leq \bar{p}^{\star j}$ which implies

[67] To save notation, the natural restriction $\omega \in [0, 1]$ is suppressed throughout.

[68] Note that $\beta = \frac{1}{2}$ will result in unrestricted $\alpha \in (0, \infty)$.

$\tau^{\star j} \leq \bar{\tau}^{\star} \equiv \frac{(2\beta+1)\alpha-1}{2+4\alpha}$, there is a technical issue in the current setting, namely, the above bliss points may not be attainable if $\bar{\tau}^{\star j}_{po} > \bar{\tau}^{\star}$.[69] To handle this problem, I now define the (constrained) bliss point policies as $\tau^{\star 1}_{po} \equiv \min(\bar{\tau}^{\star}, \bar{\tau}^{\star 1}_{po})$, $\tau^{\star 2}_{po} \equiv \min(\bar{\tau}^{\star}, \bar{\tau}^{\star 2}_{po})$, which specifies that the bliss point $\tau^{\star j}_{po}$ will be set to $\bar{\tau}^{\star}$ whenever $\bar{\tau}^{\star j}_{po} > \bar{\tau}^{\star}$. And, unless explicitly expressed, I will refer to $(\tau^{\star 1}_{po}, \tau^{\star 2}_{po})$ as bliss points from now on.

The constraint of reciprocity requires that $p^w \equiv \frac{2\beta-\alpha\tau^h-\tau^{\star 1}-\tau^{\star 2}}{2+\alpha} = p^w_0 \equiv \frac{2\beta+1+(2\beta-1)\alpha}{2+4\alpha}$, implying $\tau^{\star 1} + \tau^{\star 2} = \frac{-[(2\beta-1)\alpha-2](\alpha-1)}{2+4\alpha} - \alpha\tau^h \equiv \pi^h$. Thus, the concession jointly preferred by $\star 1$ and $\star 2$, $\tau^{\star 1}_{po} + \tau^{\star 2}_{po}$ can be less than, equal to or larger than π^h.[70]

B. Home country's offer

This subsection solves the home country's decision problem.

Lemma 4. *Assume, without loss of generality, $\star 1$ is the principal supplier, that is $\omega > \frac{1}{2}$, then there exists $0 < \alpha_0 < \alpha_1 < \alpha_2 < \alpha_3$ such that, conditional on $\alpha_i \leq \frac{2}{2\beta-1}$, $i \in \{0,1,2,3\}$, (1) for any $\alpha \in (0, \alpha_0)$, $\tau^h = 0$ (home country's bliss point) is accepted by the foreign countries of all types. (2) for any $\alpha \in (\alpha_0, \alpha_1)$, $\tau^h = 0$ is accepted only if at least one of the foreign exporters is of high type. (3) for any $\alpha \in (\alpha_1, \alpha_2)$, $\tau^h = 0$ is accepted only if the principal supplier is of high type. (4) for any $\alpha \in (\alpha_2, \alpha_3)$, $\tau^h = 0$ is accepted only if both of the foreign exporters are of high type. (5) for any $\alpha > \alpha_3$, $\tau^h = 0$ is rejected by the foreign countries of any types.*

[69] Specifically, countries' payoff functions will be *linear* when the local price exceeds the choke price above which domestic demand is driven to zero, creating a kink in the payoff function. Thus the (interior) bliss point previously characterized by a first order condition becomes invalid whenever it results in a local price larger than the choke price. However, this issue will not be present under a more general demand/supply system.

[70] Bagwell and Staiger (forthcoming) also analyze a framework with these possibilities.

Proof. The home country's maximization problem can be written as:

$$\max_{\tau^h} \Pr(\pi^h) \cdot W^h \left(p^h, p_0^w\right) + (1 - \Pr(\pi^h)) \cdot W^h \left(p_0^h, p_0^w\right)$$

$$s.t. \tau^{\star 1} + \tau^{\star 2} = \pi^h$$

$$\Pr(\pi^h) = \mathbf{1}(W^{\star j} \left(p^{\star j}, p_0^w\right) \geq W^{\star j} \left(p_0^{\star j}, p_0^w\right), \, j \in \{1, 2\})$$

Given $\star 1$ and $\star 2$'s existing policy $(\tau_0^{\star 1}, \tau_0^{\star 2})$, the sum of them is

$$\pi_0 \equiv \frac{(2\beta+1)\alpha - 1 - (2+4\alpha)\omega}{(2+4\alpha)(2+\alpha)} + \frac{(2\beta+1)\alpha - 1 - (2+4\alpha)(1-\omega)}{(2+4\alpha)(2+\alpha)} = \frac{(2\beta-1)\alpha - 2}{(1+2\alpha)(2+\alpha)}$$

Note that $\tau^h = 0$ will be accepted whenever $\bar{\tau}^\star \leq 0$, that is, whenever $\alpha \leq \frac{1}{2\beta+1}$. Thus I focus on the cases where $\alpha > \frac{1}{2\beta+1}$. For convenience, I use $(\gamma^{\star 1}, \gamma^{\star 2})$ to denote that $\star j$ is of type $\gamma^{\star j}$ with $\gamma^{\star j} \in \{l, h\}$ and $j = 1, 2$. Once the foreign countries start to negotiate, they will agree to reciprocate the home country, as long as $\tau^{\star 1} + \tau^{\star 2} \geq \pi_0$ and $\tau^{\star 1} + \tau^{\star 2} \leq \min(\bar{\tau}^\star, 2\bar{\tau}_{po}^{\star 1} - \tau_0^{\star 1}) + \min(\bar{\tau}^\star, 2\bar{\tau}_{po}^{\star 2} - \tau_0^{\star 2})$, that is, if there exists any gain from trade/negotiation. Since $\tau^{\star 1} + \tau^{\star 2} = \pi^h$ under reciprocity, π^h is accepted if and only if (1) $\pi^h \geq \pi_0$ and (2) $\pi^h \leq \min(\bar{\tau}^\star, 2\bar{\tau}_{po}^{\star 1} - \tau_0^{\star 1}) + \min(\bar{\tau}^\star, 2\bar{\tau}_{po}^{\star 2} - \tau_0^{\star 2})$. Note that (1) is automatically satisfied as $\pi^h - \pi_0 = \frac{-[(2\beta-1)\alpha-2]\alpha(1+\alpha)}{(2+4\alpha)(2+\alpha)} - \alpha\tau^h > 0 \Leftrightarrow \tau^h < \tau_0^h$, that is, any agreement entails liberalization. While (1) implies an agreement of liberalization is in the direction of potential welfare improvement for the two foreign countries, (2) requires that, while liberalization is good, there can not be too much. In particular, (2) is satisfied if and only if all the following conditions hold: (2a) $\pi^h - 2\bar{\tau}^\star = -\alpha p_0^w - \alpha\tau^h = -\alpha p_0^h < 0$. (2b) $\pi^h \leq \bar{\tau}^\star + 2\bar{\tau}_{po}^{\star j} - \tau_0^{\star j}$, for $j \in \{1, 2\}$. (2c) $\pi^h \leq 2(\bar{\tau}_{po}^{\star 1} + \bar{\tau}_{po}^{\star 2}) - \pi_0$. Note that (2a) automatically holds. For (2b), it suffices to show that $\pi^h \leq \bar{\tau}^\star - \tau_0^{\star 2}$ for $\omega = 1 - \frac{(2\beta+1)\alpha-1}{2+4\alpha}$. In particular, it can be shown that $\pi^h - \bar{\tau}^\star + \tau_0^{\star 2} = -\frac{(2\beta-1)\alpha^2+1}{2+4\alpha} < 0$. Thus, (2) is satisfied if and only if (2c) holds.

As the unconstrained optimum for the home country is $\tau^h = 0$, I discuss the conditions under which this optimum can be supported. (a) $\tau^h = 0$ will be accepted by (l,l) iff $\pi^h + \pi_0 = \frac{-[(2\beta-1)\alpha-2][\alpha(1+\alpha)-4]}{(2+4\alpha)(2+\alpha)} \leq 2(\bar{\tau}_{po}^{\star 1} + \bar{\tau}_{po}^{\star 2}) = 0$. In particular, iff $\alpha \leq \alpha_0 \equiv \frac{\sqrt{17}-1}{2}$, then $\tau^h = 0$ will be accepted regardless of $\star 1$'s and $\star 2$'s political pressure. (b) $\tau^h = 0$ will be accepted by (l,h) iff $\pi^h + \pi_0 = \frac{-[(2\beta-1)\alpha-2][\alpha(1+\alpha)-4]}{(2+4\alpha)(2+\alpha)} \leq 2(h-1)(1-\omega)$, which pins down the cutoff α_1. (c) $\tau^h = 0$ will be accepted by (h,l) iff $\pi^h + \pi_0 = \frac{-[(2\beta-1)\alpha-2][\alpha(1+\alpha)-4]}{(2+4\alpha)(2+\alpha)} \leq 2(h-1)\omega$, which pins down the cutoff α_2. (d) $\tau^h = 0$ will be accepted by (h,h) iff $\pi^h + \pi_0 = \frac{-[(2\beta-1)\alpha-2][\alpha(1+\alpha)-4]}{(2+4\alpha)(2+\alpha)} \leq 2(h-1)$, which pins down the cutoff α_3. Because the function $\frac{-[(2\beta-1)\alpha-2][\alpha(1+\alpha)-4]}{(2+4\alpha)(2+\alpha)}$ is increasing in $\alpha > 0$, I must have $0 < \alpha_0 < \alpha_1 < \alpha_2 < \alpha_3$. □

Corollary 3. *Suppose the home country does not care about delay, that is it has a discount factor of* 1*, which can make a take-it-or-leave-it offer, then in any agreement* $(\tau^h, \tau^{\star 1}, \tau^{\star 2})$ *under MFN and Reciprocity, a sufficient condition for* $\tau^h = 0$ *is* $\alpha \leq \min(\frac{\sqrt{17}-1}{2}, \frac{2}{2\beta-1})$.

C. Proofs in Section 5

This subsection solves for the bargaining equilibrium, and gives proofs for the lemmas and propositions.

Lemma 1.

Proof. As is standard, the complete information equilibrium will be the basis for my incomplete information game. In the complete information case, it is a Rubinstein model in the sense that countries will either compete to make concessions or they will compete to avoid concession.[71] Since the time between offers can be arbitrarily small, the complete information bargaining is equivalent to a Nash bargaining.[72] In particular, I denote g_{ij}, which characterizes the gap between the foreign countries' local prices $(p^{\star 2} - p^{\star 1})$,[73] as the complete information equilibrium outcome when $\star 1$

[71]However, with private information on the political pressure, whether countries compete to make or avoid concession might be uncertain to either or both of the countries.
[72]See Binmore, Rubinstein and Wolinsky (1987) for this result.
[73]Or equivalently, under reciprocity, it is the gap between their export policy $(\tau^{\star 2} - \tau^{\star 1})$.

is of type i and $\star 2$ is of type j, where $i, j \in \{h, l\}$. Then g_{ij} can be characterized by the following Nash program:

$$g_{ij} = \arg\max_{g} \{[W^{\star 1}(g,i) - W_0^{\star 1}(i)]\} \cdot \{[W^{\star 2}(g,j) - W_0^{\star 2}(j)]\}$$
$$\text{s.t. } W^{\star 1}(g,i) - W_0^{\star 1}(i) \geq 0 \text{ and } W^{\star 2}(g,j) - W_0^{\star 2}(j) \geq 0$$

that is, g_{ij} characterizes the best outcome that a country could get in a separating equilibrium, which I will focus on, as anything better than g_{ij} will be rejected by the other country once their political pressures are revealed.

My proof is similar to that in Harstad (2007). I consider $\star 2$'s strategy first.[74]

(a) Suppose $\star 1$ is revealed to be of type h by making an offer at $t_{h\cdot}$. On the one hand, as a l-type $\star 2$ will not be able to convince $\star 1$ that it is of h-type,[75] a l-type $\star 2$ will accept any $g \geq g_{hl}$, and reject any $g < g_{hl}$ and counteroffer $g = g_{hl}$ immediately. Thus, the timing, $t_{h\cdot}$ of a h-type $\star 1$'s offer is effectively the time of acceptance by a l-type $\star 2$, which can be simply rewritten as t_{hl}. On the other hand, if a h-type $\star 2$ convinces $\star 1$ that it is of type h, then $\star 1$ will accept any offer $g \leq g_{hh}$. Thus $\star 2$ will maximize its payoff under the incentive compatibility constraint by which it could truthfully signal its high political pressure :

$$\max_{d, t_{hh}} \frac{1}{\ln(\delta^{-1})} \{\delta^{t_{hh}} [W^{\star 2}(g,h) - W_0^{\star 2}(h)] + W_0^{\star 2}(h)\}$$
$$\text{s.t. } \delta^{t_{hl}} [W^{\star 2}(g_{hl}, l) - W_0^{\star 2}(l)] \geq \delta^{t_{hh}} [W^{\star 2}(g,l) - W_0^{\star 2}(l)]$$
$$g \leq g_{hh}$$

This yields $g^\star = g_{hh}$ and

$$\delta^{t_{hh} - t_{hl}} = \min\{\frac{W^{\star 2}(g_{hl}, l) - W_0^{\star 2}(l)}{W^{\star 2}(g_{hh}, l) - W_0^{\star 2}(l)}, 1\}$$

[74] If $\star 2$ proposes first, then I can change subscriptij to ji, and superscript $\star i$ to $\star j$ in the d functions and t functions.

[75] Because h will not mimic l, the intuitive criterion implies the l will respond immediately.

Note that if $g_{po}^{\star 2}(l) - g_{hl} \leq |g_{po}^{\star 2}(l) - g_{hh}|$, where $g_{po}^{2}(l) = 2(l-1)(1-\omega) = 0$, then $t_{hh} - t_{hl} = 0$, that is, the counteroffer g_{hh} itself is sufficient to signal $\star 2$'s type.

(b) Suppose $\star 1$ is revealed to be of type l by making an offer at $t_{l\cdot}$. Then similar to my argument in (a), a l-type $\star 2$ will only accept $g = g_{ll}$, reject anything different from g_{ll} and counteroffer g_{ll} immediately, where $t_{l\cdot}$ can be rewritten as t_{ll}. And a type-h $\star 2$ will maximize:

$$\max_{d, t_{lh}} \frac{1}{\ln(\delta^{-1})} \{\delta^{t_{lh}}[W^{\star 2}(g,h) - W_0^{\star 2}(h)] + W_0^{\star 2}(h)\}$$
$$\text{s.t. } \delta^{t_{ll}}[W^{\star 2}(g_{ll}, l) - W_0^{\star 2}(l)] \geq \delta^{t_{lh}}[W^{\star 2}(g,l) - W_0^{\star 2}(l)]$$
$$g \leq g_{lh}$$

This yields $g^\star = g_{lh}$ and

$$\delta^{t_{lh} - t_{ll}} = \min\{\frac{W^{\star 2}(g_{ll}, l) - W_0^{\star 2}(l)}{W^{\star 2}(g_{lh}, l) - W_0^{\star 2}(l)}, 1\}$$

Since $g_{ll} = 0$, a $\star 2$ of type l will not mimic a type h whatsoever.[76] Thus $t_{lh} - t_{ll} = 0$. This feature is due to the existence of bliss points in the range of feasible agreements when countries are of low type, which results in the slackness of the incentive compatibility condition.

I turn to $\star 1$'s strategy next.

(c) If $\star 1$ is of type l, it will not be able to convince $\star 2$ that it is of type h in a separating equilibrium[77]. As a result, it will make an offer that is acceptable to a $\star 2$ of type l, that is, $g = g_{ll}$ at $t_{ll} = 0$. Thus $\star 1$'s expected payoff will be

$$\bar{W}^{\star 1} \equiv (1-p)\frac{1}{\ln(\delta^{-1})} W^{\star 1}(g_{ll}, l) + p\frac{1}{\ln(\delta^{-1})}\{\delta^{t_{lh}}[W^{\star 1}(g_{lh}, l) - W_0^{\star 1}(l)] + W_0^{\star 1}(l)\}$$

[76] $g = 0$ is also the best offer for a country of type l. This will not be true in cases where $l > 1$.

[77] In my setting, $\star 1$ can not make a pooling offer by choosing either g_{lh} or g_{ll}, as both of them can screen $\star 2$ perfectly.

(d) If $\star 1$ is of type h, it will maximize its expected payoff subject to the incentive compatibility constraint that the offer should be unattractive to a $\star 1$ of type l, but acceptable for a $\star 2$ of type l:

$$\max_{d,t_{hl}} \left\{ \begin{array}{c} p\frac{1}{\ln(\delta^{-1})}\{\delta^{t_{hh}}[W^{\star 1}(g_{hh},h) - W_0^{\star 1}(h)] + W_0^{\star 1}(h)\} \\ +(1-p)\frac{1}{\ln(\delta^{-1})}\{\delta^{t_{hl}}[W^{\star 1}(g,h) - W_0^{\star 1}(h)] + W_0^{\star 1}(h)\} \end{array} \right\}$$

$$\text{s.t. } \bar{W}^{\star 1} \geq p\frac{1}{\ln(\delta^{-1})}\{\delta^{t_{hh}}[W^{\star 1}(g_{hh},l) - W_0^{\star 1}(l)] + W_0^{\star 1}(l)\}$$
$$+ (1-p)\frac{1}{\ln(\delta^{-1})}\{\delta^{t_{hl}}[W^{\star 1}(g,l) - W_0^{\star 1}(l)] + W_0^{\star 1}(l)\}$$
$$g \geq g_{hl}$$

This yields

$$\delta^{t_{hl}} = \min\{\frac{p\delta^{t_{lh}}[W^{\star 1}(g_{lh},l) - W_0^{\star 1}(l)] + (1-p)[W^{\star 1}(g_{ll},l) - W_0^{\star 1}(l)]}{p\delta^{t_{hh}-t_{hl}}[W^{\star 1}(g_{hh},l) - W_0^{\star 1}(l)] + (1-p)[W^{\star 1}(g_{hl},l) - W_0^{\star 1}(l)]}, 1\}$$

Substituting the expressions for $W_0^{\star j}(l)$ and $W^{\star j}(g_{lh},l)$ into the above expressions for t's, I have:

$$\delta^{t_{hh}-t_{hl}} = \min\{\frac{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{hl}}{2})[-\frac{g_{hl}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{hh}}{2})[-\frac{g_{hh}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}, 1\}$$
$$\delta^{t_{lh}} = \min\{\frac{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{ll}}{2})[-\frac{g_{ll}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}{(\frac{5}{18} - \frac{1}{3}\omega + \frac{g_{lh}}{2})[-\frac{g_{lh}}{4} + \frac{5}{36} - \frac{1}{6}\omega]}, 1\} \equiv 1$$

$$\delta^{t_{hl}} = \min\{\frac{(\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{lh}}{2})[\frac{g_{lh}}{4} + \frac{1}{6}\omega - \frac{1}{36}] + (\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{ll}}{2})[\frac{g_{ll}}{4} + \frac{1}{6}\omega - \frac{1}{36}]}{\delta^{t_{hh}-t_{hl}}(\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{hh}}{2})[\frac{g_{hh}}{4} + \frac{1}{6}\omega - \frac{1}{36}] + (\frac{1}{3}\omega - \frac{1}{18} - \frac{g_{hl}}{2})[\frac{g_{hl}}{4} + \frac{1}{6}\omega - \frac{1}{36}]}, 1\}$$

□

Characterization of equilibrium agreements. The Nash solution can be characterized by the first order condition of the Nash bargaining problem[78]:

$$\frac{1}{16}g^3+\frac{3}{16}[(\gamma^{\star 1}-1)\omega-(\gamma^{\star 2}-1)(1-\omega)]g^2$$
$$+\frac{1}{36}[-(\omega^2-\omega+\frac{13}{36})-(3\omega-\frac{1}{2})(\gamma^{\star 1}-1)\omega-(\frac{5}{2}-3\omega)(\gamma^{\star 2}-1)(1-\omega)$$
$$-18(\gamma^{\star 1}-1)\omega(\gamma^{\star 2}-1)(1-\omega)]g+\frac{1}{36}[-6(\gamma^{\star 1}-1)\omega(\gamma^{\star 2}-1)(1-\omega)(1-2\omega)-$$
$$(\gamma^{\star 1}-1)\omega(\frac{5}{6}-\omega)^2+(\gamma^{\star 2}-1)(1-\omega)(\omega-\frac{1}{6})^2]=0$$

In particular, it is straightforward to show $g_{ll}=0$, and g_{hh} solves

$$\frac{1}{16}g^3+\frac{3}{16}(h-1)(2\omega-1)g^2+\frac{1}{36}[-(\omega^2-\omega+\frac{13}{36})$$
$$-(h-1)(6\omega^2-6\omega+\frac{5}{2})-18(h-1)^2\omega(1-\omega)]g$$
$$+\frac{1}{36}[-6(h-1)^2\omega(1-\omega)(1-2\omega)-(h-1)\omega(\frac{5}{6}-\omega)^2+(h-1)(1-\omega)(\omega-\frac{1}{6})^2]=0$$

g_{hl} solves

$$\frac{1}{16}g^3+\frac{3}{16}(h-1)\omega g^2+\frac{1}{36}[-(\omega^2-\omega+\frac{13}{36})-(3\omega-\frac{1}{2})(h-1)\omega]g+\frac{1}{36}[-(h-1)\omega(\frac{5}{6}-\omega)^2]=0$$

g_{lh} solves

$$\frac{1}{16}g^3-\frac{3}{16}(h-1)(1-\omega)g^2+\frac{1}{36}[-(\omega^2-\omega+\frac{13}{36})-(\frac{5}{2}-3\omega)(h-1)(1-\omega)]g$$
$$+\frac{1}{36}[(h-1)(1-\omega)(\omega-\frac{1}{6})^2]=0$$

These equations, together with second order conditions, yield the analytical solutions for $\{g_{ij}\}$, which fully characterize other variables of interest, in particular the

[78]Together with the second order condition, this polynomial yields a unique solution.

delay. For example, letting $h \to \infty$, I have :

$$\begin{aligned} g_{hl} =& -\frac{2}{9}[\sqrt{3(\frac{5}{6}-\omega)^2+(\omega-\frac{1}{6})^2}-(\omega-\frac{1}{6})] \\ g_{lh} =& \frac{2}{9}[\sqrt{3(\omega-\frac{1}{6})^2+(\frac{5}{6}-\omega)^2}-(\frac{5}{6}-\omega)] \\ g_{hh} =& \frac{2}{3}(\omega-\frac{1}{2}) \end{aligned}$$

But I currently have to rely on the numerical grids as the closed form solutions are themselves algebraically complicated, exemplified by the above example.[79] In particular, I have three parameters in the model, namely h, ω and δ where only h and ω plays an role in the bargaining delay patterns of interest. Thus I resort to the grid method that searches all possible values in the domain of parameters, where the propositions and lemmas are established in the way similar to the following.

Proposition 1.

Proof. In this parametric setting, the only parameters/variables governing the sign of the derivative $\frac{d(t_{hh}-t_{hl})}{d\omega}$ is ω and $g(h,\omega)$. In other words, this sign is fully characterized by the two parameters h and ω. Thus I numerically calculate the derivatives, exhausting the space $(h,\omega) \in (1,\bar{h}) \times [\frac{1}{6},\frac{5}{6}]$ by taking sufficiently fine grids. □

D. Home Long

This subsection discusses the case of home long.

An important observation in the paper is the focus on the case of home short, under which only the reciprocity in negotiation (going-down) plays a role, because the foreign exporting countries ($\star 1$ and $\star 2$) are competing for more concessions - nothing to renegotiate.

[79]While there exists closed-form solution, making a strict analytical proof more desirable, I don't gain much economic insight other than the algebraic complexities.

Thus, I now allow both reciprocity in negotiation and that in renegotiation, but remove the rationing rule from Bagwell and Staiger (forthcoming). Technically, proposing its own bliss point, π, is now the home country's dominant strategy and the reciprocity condition in our original paper $\tau^{\star 1} + \tau^{\star 2} = \pi$ becomes $\tau^{\star 1} + \tau^{\star 2} \leq \pi$. Further note that any agreement with $\tau^{\star 1} + \tau^{\star 2} < \pi$ cannot be an equilibrium as at least one of the foreign countries can profitably deviate. Thus our analysis in the original paper (home short) still holds.

In this model, high type country wants more concession (subsidy increase/tariff decrease) relative to low type country. Thus I define $\tau_{ij}^{\star\star} \equiv \tau_{po}^{\star 1}(i) + \tau_{po}^{\star 2}(j)$ where $\tau_{po}^{\star 1}$ and $\tau_{po}^{\star 2}$ is explicitly dependent on their respective types i and j. Note that, for a given ω, it can be ranked as $\tau_{ll}^{\star\star} < \tau_{lh}^{\star\star} < \tau_{hl}^{\star\star} < \tau_{hh}^{\star\star}$. Thus the following cases can be discussed:

(1) $\pi \leq \tau_{ll}^{\star\star}$. Delay in (pure) home short, as in the original paper.

(2) $\pi \geq \tau_{hh}^{\star\star}$. In this case, it is a dominant strategy to propose $\tau_{po}^{\star 1}(i)$ and $\tau_{po}^{\star 2}(j)$, because $\tau_{po}^{\star 1}(i) + \tau_{po}^{\star 2}(j) \leq \pi$. No delay.

(3) $\tau_{ll}^{\star\star} < \tau_{lh}^{\star\star} < \tau_{hl}^{\star\star} \leq \pi < \tau_{hh}^{\star\star}$. In this case, proposing $\tau_{po}^{\star 1}(i)$ is a dominant strategy for low type $\star 1$. Thus a high type $\star 1$ can signal its type by simply proposing its bliss point $\tau_{po}^{\star 1}(h)$ which a low type $\star 2$ accepts immediately, while a high type $\star 2$ counteroffers g_{hh} immediately which is accepted by $\star 1$ immediately. No delay.

(4) $\tau_{ll}^{\star\star} < \tau_{lh}^{\star\star} \leq \pi < \tau_{hl}^{\star\star} < \tau_{hh}^{\star\star}$. In this case, proposing $\tau_{po}^{\star 1}(i)$ is still a dominant strategy for low type $\star 1$. Thus a high type $\star 1$ can signal its type by simply proposing its bliss point $\tau_{po}^{\star 1}(h)$. However, now the low type $\star 2$ has the incentive to mimic a high type if $g_{hh} \overset{2}{\succ} g_{hl}$. Delay possible when high type $\star 2$ signals its type.

(5) $\tau_{ll}^{\star\star} < \pi < \tau_{lh}^{\star\star} < \tau_{hl}^{\star\star} < \tau_{hh}^{\star\star}$. In this case, if $\star 1$ is revealed to be a high type, then the low type $\star 2$ has the incentive to mimic a high type if $g_{hh} \overset{2}{\succ} g_{hl}$. if $\star 1$ is revealed to be low type, then low type $\star 2$ will not mimic. This is the same case as case 1, as we assume the low type to be free trader – each free trader will do free trade in equilibrium regardless of its export share.

Chapter 2. A Lid on WTO Retaliation

9. Introduction

The question why trade agreements, or more specifically the GATT/WTO, are necessary, has been studied extensively where either the terms-of-trade externality approach or the commitment approach among others has been applied. However, another very important issue regarding the WTO/GATT framework concerns how countries may self-enforce those trade agreements, as is pointed out by Bagwell & Staiger (2002): "*...it is often observed that the pillars of GATT are the principles of reciprocity and nondiscrimination (MFN), while enforcement mechanisms form the heart of the GATT system...*". In particular, how could countries possibly achieve and sustain cooperation in the international environment where no counterpart of the relatively complete domestic legal system exists? The classical approach builds on the structure of infinitely-repeated games. For example, the analysis by Dixit (1987) shows that countries choose to cooperate because they care about their continuation payoff in an infinitely-repeated game. However, directly applying this classical setting to tariff games seems unable to account for the on-equilibrium-path deviations from cooperation, which would appear to be relatively common in the GATT/WTO practice: the stationarity of the equilibria predicts that countries either cooperate or stay at the static Nash equilibrium forever on the equilibrium path. A natural way to generate those deviations is to introduce uncertainty facing each decision-making country.[80] Following the more recent imperfect private monitoring structure in repeated games, Park (2011) proposes that observability issues, due to the concealed trade barriers, can also lead to failure/inefficiency of coordination among trading partners. More interestingly, he argues that the presence of the WTO may restore

[80] An alternative is to interpret WTO disputes not as reflecting on-equilibrium path deviations from cooperation, but rather as on-equilibrium path rebalancing of incentive constraints (as in Bagwell and Staiger, 1990) or on-equilibrium path opportunism within the bounds permitted under incomplete contracts (as in Maggi and Staiger, 2011).

the cooperation in a more efficient way, via comparing "Private Triggering Strategy (PTS)" with "Third-party Triggering Strategy (TTS)". Nonetheless, in practice, the WTO does not allow for arbitrary retaliations once the punishment is triggered. In particular, the principle of reciprocity plays a role, which means there is an upper bound – a "lid" – for the retaliatory tariff rate. As a result, this restriction tends to reduce the pain from punishment, possibly making cooperation even harder.

The main objective of this paper is to examine the rationale for this lid feature, in the context of imperfect private monitoring.[81] In particular, I aim to answer the question why the WTO wants to put such restrictions from the perspective of welfare improvement. The key idea is that when the WTO has more instruments, specifically the lid, it may help improve the welfare because different instruments result in distinctive welfare cost. Thus, if the WTO were to impose a retaliation phase so as to enhance cooperation, it might be optimal to choose one tool over the others depending on the environment. A further goal of this paper is to examine several different mechanism to sustain such restrictions, and to do so I need to consider the verifiability of countries' concealed actions. Thus several approaches are discussed, among which the discount factor and the fixed cost of establishing an investigatory "Panel" determines whether the lid helps countries improve their welfare and, if so, whether countries call upon the Panel at all time.

I focus on an environment where symmetric countries cannot observe or precisely infer each other's actions, i.e. the total protection level in a two-country model. This protection-setting game is repeated infinitely, where the *ex post* realized payoff is each country's private signal, and the explicit tariff is publicly observable to both countries.

[81]Repeated games with imperfect private monitoring have not been well understood in general, compared to the public monitoring cases, see Kandori (2002) for a good survey. While communication is a classical tool as in Kandori & Matsushima (1998), Compte (1998) among others, Matsushima (2004), Yamamoto (2007), Sugaya (2011) provide more general analysis without communication. More recently, Cherry & Smith (2012) examines the tightest upper bound for the possible sequential equilibrium payoffs in this class of games, via introducing the Markov Perfect Correlated Equilibrium.

To make the results comparable, the model adopts the same primitives as in Park (2011). Although countries follow the WTO's command regarding the triggering time, a vital issue arises that it should be incentive-compatible for countries to comply with the WTO's command, i.e. the retaliation lid, during the punishment phase, because countries can always attain the static Nash punishment via applying the concealed trade barriers. After showing that the WTO can do strictly better via the lid, I next introduce the fixed cost of establishing a Panel into the model, to examine how this extension affects countries' decision regarding the establishment of a Panel in different phases. It is shown that depending on the discount factor, as long as the fixed cost is above some cut-off value, there is always an incentive for countries to request a Panel only when a dispute actually occurs, i.e. when the retaliation is triggered.

The paper is organized as follows: the basic framework is presented in Section 10. The comparative analysis is conducted in Section 11. Section 12 augments the model with costly establishment of the WTO Panel in its dispute settlement procedure, taking the incentive of compliance into consideration. In Section 13, I assume the concealed trade barriers and explicit tariff are imperfect substitutes, as opposed to the case in Park (2011) where the substitution is perfect. Section 14 concludes the paper.

10. The Model

This section consists of the basic elements of the repeated protection-setting game, as introduced by Park (2011).[82] In the 2-by-2 trade model, the home country imports good 2 and exports good 1, while the foreign country (F) imports good 1 and exports good 2. For simplicity, I assume that the countries are symmetric, choosing their total protection levels and explicit tariffs, respectively denoted by $(\tau, \tau^{\star})$and $(e, e^{\star})$,

[82]All the notations follow Park (2011), unless otherwise noted.

in an infinitely repeated game.[83] And I assume that the interaction between two countries only travels through international relative price. Denote the (local) price vectors in H and F as (p_1, p_2) and $(p_1^\star, p_2^\star)$, respectively. Thus, the following equations hold: $p_2 = (1+\tau)\, p_2^\star$, $p_1^\star = (1+\tau^\star)\, p_1$.[84] Define the international relative price, or equivalently F's terms of trade, as $\pi \equiv \frac{p_2^\star}{p_1}$. The main feature of the model is that each country's total protection level cannot be observed by the other one. In other words, $\tau - e$, interpreted as the "concealed trade barrier", is only observable to H, and similarly $\tau^\star - e^\star$ is only observable to F.[85] Regarding the source of uncertainty – so that on-equilibrium path deviation is possible, there are three fundamental shocks that enter into each country's expectation: $\theta, \theta^\star, \epsilon^u$,[86] which respectively corresponds to the domestic shock privately observed by the home country in the end of each period , the foreign shock privately observed by the foreign country in the end of each period, and the common shock observable to neither country in any time.[87] The balanced trade condition $\pi \cdot m\,(\pi, \tau, \theta, \epsilon^u) = m^\star\,(\pi, \tau^\star, \theta^\star, \epsilon^u)$ pins down π, where m and $m^\star$ is respectively the import demand function of H and F. Because of the symmetry, I simply assume that H's (expected) one-period payoff can be written as $u\,(\tau, \tau^\star)$, satisfying $\frac{\partial u}{\partial \tau}|_{\tau=0} > 0$, $\frac{\partial u}{\partial \tau^\star} < 0$, $\frac{\partial^2 u}{\partial \tau^2} < 0$, $\frac{\partial^2 u}{\partial \tau \partial \tau^\star} = 0$, and $\frac{\partial (u+u^\star)}{\partial \tau} < 0$. And F's payoff is similarly defined as $u^\star\,(\tau^\star, \tau)$. In particular, these conditions imply a dominant strategy for each country, which I denote as h, determined by $\frac{\partial u}{\partial \tau}|_{\tau=h} = 0$. Moreover,

[83]Following the convention in the literature, a superscript of $\star$ always denotes the foreign country.

[84]Thus, the explicit tariff e ($e^\star$) does not affect countries' payoff, instead it serves as a coordination device/signal. Only the total protection level τ ($\tau^\star$) matters for the payoff.

[85]I assume that each country cannot infer the other country's total protection level via the arbitrage condition that links the domestic price and international price. One possible justification is market segmentation, which essentially breaks that link. However, I can also assume that prices are noisy, or at least the observations of them are noisy. For example, $p_1 = \theta\,(1+\tau)\, p_2^\star$, where θ is a stochastic term which also breaks the link between the (observed) $p_1$and $(1+\tau)\, p_2^\star$. Actually, this formulation is similar to the one in Green and Porter (1984), where the demand was subject to some random shocks.

[86]If there is no shocks, countries can coordinate via their payoffs equally well as in cases with perfect information. The noise mentioned in footnote 85 can also be incorporated into the θ here.

[87]ϵ^u serves to enabling the comparison of welfare during the WTO era and the non-WTO era later.

as a benchmark, free trade is the first-best outcome in this setting. These assumptions essentially make the framework a special case of the analysis by Dixit (1987) in terms of the specification of each country's best response.[88] At the end of every period, each country can observe a signal $\omega_t^i \equiv (u_t^i, \theta_t^i) \in \Omega \subseteq \mathbb{R}^2$, with $i \in \{ , \star\}$, where u_t^i is the *realized* payoff at period t.[89] Based on those signals, they choose the actions in next period(s). Note that each action profile is a tuple $(a, a^\star) \equiv ((\tau, e), (\tau^\star, e^\star)) \in A \times A^\star$, where A and $A^\star$ are also subsets of $\mathbb{R}^2$. Thus, the strategy profile can be written as $\left((s_t)_{t\geqslant 1}, (s_t^\star)_{t\geqslant 1}\right)$, with $s_t : A^{t-1} \times \Omega^{t-1} \times (E^\star)^{t-1} \to A$, where a superscript $t-1$ denotes the history before period t, which is known to the decision maker prior to the beginning of period t. And $(s_t^\star)_{t\geqslant 1}$ can be specified symmetrically. In addition, countries discount the future payoff at a common rate $\delta^C \in (0,1)$.

On the one hand, if there is no WTO, upon observing their private signals, countries behave non-cooperatively and it is hard to analyze the complete set of strategy profiles. Park (2011) analyzes the set of Private Triggering Strategies (PTS), in which retaliations are triggered as long as $\omega_t^i \in \Omega^D \subseteq \Omega^i$, where Ω^D is the triggering set.[90] On the other hand, with the WTO in the model, Park (2011) also analyzes the set of Third-party Triggering Strategies (TTS), in which countries report the signals they receive to the WTO, and follow the WTO's advice on their actions.[91] Formally, under the WTO era, countries' strategy can be written as $\left((s_t)_{t\geqslant 1}, (s_t^\star)_{t\geqslant 1}\right)$, with $s_t : A^{t-1} \times (E^\star)^{t-1} \times M^{t-1} \times (M^\star)^{t-2} \to A$,[92] where there is a new component of the

[88]See Bagwell & Staiger (2002) for a good summary.

[89]As pointed out in Riezman (1991), the set of signals upon which strategies are based is critical, in that if the triggering is based upon the international relative price, there exists no equilibrium. This is in sharp contrast to Green and Porter (1984), where the triggering can be based on observed market price. I simply assume countries do not use π as a triggering signal.

[90]Ω^D can be chosen to maximize welfare as long as it can be supported as an equilibrium.

[91]While the mechanism that enables countries to truthfully report their signals is assumed to be outside of the model, in any equilibrium, incentive constraints are imposed so as to make sure countries follow the WTO's advice.

[92]F's strategy can be specified symmetrically. Again, a superscript $t-1$ denotes the history before period t, which is known to the decision maker prior to the beginning of period.

history M^i: the advice (or "Ruling" in the WTO terminology) made to country i by the WTO. And the additional assumption is that the WTO's advice to one country only becomes known to the other country after one period.[93] Since I am interested in comparing the case under the WTO without retaliation restrictions – the one in Park (2011), and the case under the WTO with retaliation restrictions considered in this paper, focusing on the following Third-party Triggering Strategies (TTS) suffice for this purpose.

(1) if $t-1$ is *cooperative*, i.e., $\left(e_{t-1}, e^{\star}_{t-1}\right) = (0,0)$, each country plays $(l,0)$ unless told by the WTO to punish.

(2) if $t-1$ is *cooperative*, the WTO tells H to play $(\bar{\tau},\bar{\tau})$, if $\omega_{t-1} \in \Omega^D$ and tells F to play $(\bar{\tau},\bar{\tau})$, if $\omega^{\star}_{t-1} \in \Omega^D$.[94]

(3) If a *punishment* phase is initiated in period $t-1$, $\left(e_{t-1}, e^{\star}_{t-1}\right) \neq (0,0)$, then: (a) if $e_{t-1} \cdot e^{\star}_{t-1} = 0$, $(\bar{\tau},\bar{\tau})$ is played for the next $T-2$ periods and continues to play $(\bar{\tau},\bar{\tau})$ for one more period with probability λ.[95] (b) if $e_{t-1} \cdot e^{\star}_{t-1} \neq 0$, $(\bar{\tau},\bar{\tau})$ is played for the next $T^S - 2$ periods and continues to play $(\bar{\tau},\bar{\tau})$ for one more period with probability λ^S.

(4) In period 1 and other *initial* periods that start directly after the end of any punishment phase, H plays $(l,0)$ with probability $1 - Pr(l)$, but initiates a

[93]As is pointed out in by Park (2011), this assumption only serves to enabling the comparability of welfare under different eras.

[94]In Park (2011), $(\bar{\tau},0)$ is used instead, which looks like a typo. However, because the third-party, the WTO knows whether a punishment has been triggered and the explicit protection e does not enter into countries' payoff function, this does not affect the calculation.

[95]This λ is from a public "device", and so is the λ^s in the simultaneous-deviation case below. This serves to smooth the payoffs.

punishment phase by playing $(\bar{\tau}, \bar{\tau})$ with $Pr(l)$ defined as,[96]

$$Pr(l) \equiv Pr\left(\omega_t \in \Omega^D | (l,0), (l,0)\right)$$

Then, I denote the expected discounted life-time payoff for H at period 0 or any initial periods as V, given those strategy specification, where

$$\begin{aligned}
V = & (1-Pr)^2 \left(u(l,l) + \delta^C V\right) \\
& + Pr(1-Pr)\left(u(l,\bar{\tau}) + \frac{\delta^C - {\delta^C}^{T-1}}{1-\delta^C} u(\bar{\tau},\bar{\tau}) + {\delta^C}^{T-1}\lambda u(\bar{\tau},\bar{\tau}) + \delta V\right) \\
& + Pr(1-Pr)\left(u(\bar{\tau},l) + \frac{\delta^C - {\delta^C}^{T-1}}{1-\delta^C} u(\bar{\tau},\bar{\tau}) + {\delta^C}^{T-1}\lambda u(\bar{\tau},\bar{\tau}) + \delta V\right) \\
& + Pr^2\left(u(\bar{\tau},\bar{\tau}) + \frac{\delta^C - {\delta^C}^{T^S-1}}{1-\delta^C} u(\bar{\tau},\bar{\tau}) + {\delta^C}^{T^S-1}\lambda^S u(\bar{\tau},\bar{\tau}) + \delta^S V\right)
\end{aligned} \tag{10.1}$$

where $\delta \equiv \lambda {\delta^C}^T + (1-\lambda){\delta^C}^{T-1}$, $\delta^S \triangleq \lambda^S {\delta^C}^{T^S} + \left(1-\lambda^S\right){\delta^C}^{T^S-1}$

If $\bar{\tau} = h$, the TTS here coincides with that in Park(2011). Define $T^W \in [1,\infty)$ such that $\delta^C - \left(\delta^C\right)^{T^W} = \delta^C - \delta$.[97] Following Park (2011), I assume that $\delta^C - \delta^S = 2\left(\delta^C - \delta\right)$.[98] It follows that

$$V = \frac{[1-Pr(l)][u(l,l) - u(h,h)]}{1-\delta^C + 2Pr(l)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]} + \frac{u(h,h)}{1-\delta^C} \tag{10.2}$$

The third party (the WTO in this context) chooses T^W and the parameter that controls Pr to maximize the above payoff function. However, in a typical dispute case,

[96]As a typical specification in repeated games with imperfect private monitoring, I also assume the full support condition holds: $Pr\left(\omega_t, \omega_t^\star | (a, a^\star)\right) > 0$ for any $(\omega_t, \omega_t^\star, a, a^\star) \in \Omega \times \Omega^\star \times A \times A^\star$, and conditional independence holds between $\omega_t, \omega_t^\star$.

[97]Note that $T^W = 1$ implies $T = 2$ and $\lambda = 0$.

[98]Actually, this condition is necessary for countries to truthfully represent its private signals in the PTS case, i.e. when there is no WTO. But here the condition is assumed to hold so as to enable the comparison between my result and Park (2011). Note that if this condition is relaxed, there are even less constraints, implying a larger welfare gain.

countries are only authorized by the WTO to retaliate up to a certain level, which I denote as $\bar{\tau}$ in the above-specified TTS, instead of the duration of retaliation. This seems puzzling at first glance as countries are constrained in their ability to retaliate against the deviation from cooperation – thus incentives to cooperate become smaller. In order to resolve this "puzzle", I now proceed to the comparison of the welfare with the retaliation lid and the welfare without it in the next section, to examine the possible roles this restriction plays.

11. Perfect Verifiability of Retaliatory Protection

11.1. **Upper Bound.** Suppose the way that the social planner controls $Pr\left(l\right)$ is via the so-called *trigger control* ω^D, satisfying $\frac{\partial^2 Pr(l)}{\partial\tau\partial\omega^D} > 0$ and $\frac{\partial Pr(l)}{\partial\omega^D} > 0$. Specifically, I interpret $\frac{\partial Pr(l)}{\partial\tau}$ as the *sensitivity* of detecting potential deviation, and interpret $1 - Pr\left(l\right)$ as the *stability* of detecting potential deviations. Thus, the previous conditions imply that the social planner is facing a trade-off between *sensitivity* and *stability*: increasing the *trigger control* ω^D results in higher sensitivity, but it also leads to lower stability. In particular, the optimization problem facing the third party (social planner) can be written as:[99]

[99]The condition $I^W\left(l\right)$ is the incentive constraint for countries to implement l in the cooperative phase, resulted from the dynamic programming problem facing the countries. See Park (2011) for more details.

$$
\begin{aligned}
&\max_{\omega^D, T^W, \bar{\tau}} \left\{ \frac{[1 - Pr(l)][u(l,l) - u(\bar{\tau},\bar{\tau})]}{1 - \delta^C + 2Pr(l)\left[\delta^C - (\delta^C)^{T^W}\right]} + V_N^W \right\} \\
&\text{s.t.} \quad I^W(l) = 0 \\
&\text{where} \quad I^W(l) \equiv \frac{\partial u(l,l)}{\partial \tau} - \delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)] \times \\
&\qquad \left\{ u(l,l) - u(l,\bar{\tau}) + \left(\delta^C - \left(\delta^C\right)^{T^W}\right)\left(V_C^W - V_N^W\right) \right\} \\
&\text{and} \quad V_N^W \equiv \frac{u(\bar{\tau},\bar{\tau})}{1 - \delta^C} \\
&\text{and} \quad V_C^W \equiv \frac{(1 - Pr(l))[u(l,l) - u(\bar{\tau},\bar{\tau})]}{1 - \delta^C + 2Pr(l)\left[\delta^C - (\delta^C)^{T^W}\right]} + V_N^W
\end{aligned} \tag{11.1}
$$

Park (2011) establishes that the WTO can help improve efficiency under certain plausible conditions when $\bar{\tau} = h$. However, it is unclear whether introducing retaliation restrictions can enhance or weaken the WTO's role in terms of efficiency enhancing. Nevertheless, it is trivial to simply put the restrictions on the explicit tariff setting. Because countries always tend to use concealed trade barriers to restore the dominant total protection level to h in the punishment phase, there is no need to impose any condition to induce non-deviation incentive.[100] Thus, it naturally follows that:

Proposition 5. *When the WTO directly imposes the retaliation restriction on the observable explicit tariffs, then the concealed trade barrier in the punishment phase is non-zero, that is, countries adopt a total protection level that is higher than the explicit tariff in the punishment phase. Thus the WTO with retaliation restrictions cannot do better in this case.*

[100]Note that h is the dominant strategy in the one-shot stage game, and that signals sent during punishment phase convey no information, due to the specified triggering strategies.

Proof. $u(h,\tau^\star) > u(\tau,\tau^\star)$ for any $\tau \neq h$, and $\tau^\star$.

□

Intuitively, the WTO can always choose to set $\bar{\tau} = h$, implying that under the retaliation constraint, the WTO can help to achieve a *weakly* better outcome in any case. In particular, the first-order conditions can be written as follows:[101]

$$\begin{aligned}\frac{dV_C^W}{d\omega^D} &= \frac{\partial V_C^W}{\partial l}\frac{\partial l}{\partial \omega^D} + \frac{\partial V_C^W}{\partial Pr}\frac{\partial Pr(l)}{\partial \omega^D} = 0\\ \frac{dV_C^W}{dT^W} &= \frac{\partial V_C^W}{\partial l}\frac{\partial l}{\partial T^W} + \frac{\partial V_C^W}{\partial T^W} = 0\\ \frac{dV_C^W}{d\bar{\tau}} &= \frac{\partial V_C^W}{\partial l}\frac{\partial l}{\partial \bar{\tau}} + \frac{\partial V_C^W}{\partial \bar{\tau}} = 0\end{aligned} \tag{11.2}$$

In order to analyze the choice of $\bar{\tau}$, I write down the following expressions for partial derivatives with respect to $\bar{\tau}$:

$$\begin{aligned}\frac{\partial V_C^W}{\partial \bar{\tau}} &= \left\{\frac{-(1-Pr(l))}{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]} + \frac{1}{1-\delta^C}\right\}\cdot\frac{du(\bar{\tau},\bar{\tau})}{d\bar{\tau}} < 0\\ \frac{\partial V_C^W}{\partial l}\frac{\partial l}{\partial \bar{\tau}} &= \left[-\frac{\partial V_C^W/\partial l}{\partial I^W(l)/\partial l}\right]\cdot\frac{\partial I^W(l)}{\partial \bar{\tau}} > 0\end{aligned} \tag{11.3}$$

In establishing that $\frac{du(\bar{\tau},\bar{\tau})}{d\bar{\tau}} < 0$, I use the fact that

$$\begin{aligned}\frac{du(\bar{\tau},\bar{\tau})}{d\bar{\tau}} &= \frac{\partial u(\tau,\tau^\star)}{\partial \tau}|_{(\bar{\tau},\bar{\tau})} + \frac{\partial u(\tau,\tau^\star)}{\partial \tau^\star}|_{(\bar{\tau},\bar{\tau})}\\ &= \frac{\partial u(\tau,\tau^\star)}{\partial \tau}|_{(\bar{\tau},\bar{\tau})} + \frac{\partial u^\star(\tau^\star,\tau)}{\partial \tau}|_{(\bar{\tau},\bar{\tau})} < 0\end{aligned} \tag{11.4}$$

where the second equality holds by symmetry, and the last inequality holds by assumption.

[101]In solving the optimization, the corner solutions also need to be checked.

Now I proceed to analyze how the constraint $I^W(l) = 0$ is affected by the retaliation restrictions. For the constraint $I^W(l) = 0$, it must hold that $\frac{\partial u(l,\bar{\tau})}{\partial \bar{\tau}} < 0$, and $\frac{\partial \left(V_C^W - V_N^W\right)}{\partial \bar{\tau}} \propto -\frac{\partial u(\bar{\tau},\bar{\tau})}{\partial \bar{\tau}} > 0$, implying $\frac{\partial I^W(l)}{\partial \bar{\tau}} < 0$. In particular,

$$\begin{aligned}\frac{\partial I^W(l)}{\partial \bar{\tau}}|_{\bar{\tau}=h} =& \delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)] \times \\ & \frac{\left\{1 - \left(\delta^C\right)^{T^W} + Pr(l)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]\right\}}{1 - \delta^C + 2Pr(l)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]} \cdot \frac{\partial u(h,h)}{\partial \tau^{\star}} < 0\end{aligned} \tag{11.5}$$

Note that a decrease in $\bar{\tau}$ tightens the I^W constraint, implying that l tends to increase. Intuitively, it is due to the fact that the penalty from triggering the punishment phase becomes smaller. However, it might also be that countries can switch to "less costly" retaliation via choosing other controls such as ω^D. In particular, the social planner is facing two counter-acting incentives: enhancing cooperation through larger penalty in the punishment phase *versus* lowering the welfare loss once the punishment phase is triggered, which is a standard mechanism in repeated game settings of the kind. Moreover, one additional consideration associated with the extra instrument $\bar{\tau}$ is the trade-off between T^W, ω^D, and $\bar{\tau}$, in order to attain the appropriate retaliatory protection.

On the one hand, if the WTO can only put the lid directly on explicit tariff, there is no welfare gain in this case. On the other hand, it seems unnatural to assume that the WTO can directly put those restrictions on the total protection level, which it cannot observe. However, there exists asymmetry during different phases of the WTO practice. For example, the Panel is called upon only at the "retaliation phase", which investigates the specific cases and possibly enables the WTO to access more-detailed and/or more accurate information regarding the retaliation process.[102] Since this feature is explored in the future sections, I simply assume for now that the WTO

[102] In the WTO practice, the Panel reports to the Dispute Settlement Body (DSB) which decides whether to adopt the report from investigation.

is able to directly control the retaliatory protection in any punishment phase,[103] and then the question becomes whether the WTO can put such direct restrictions. In particular, is it beneficial for the countries? Thus, I directly impose the retaliation upper-bound, $(\bar{\tau}, \bar{\tau})$ on the original problem. In doing so, I am agnostic of why the WTO can not impose the restrictions on both phases.[104]

However, a critical issue is that, if I simply assume the WTO has access to the total protection levels in the retaliation phase, the optimization becomes trivial again. In particular, I have the following proposition.

Proposition 6. *When the WTO directly imposes the retaliation restriction on the total protection levels, which I assume it can effectively and costlessly verify during any retaliation phase for whatever reason, the WTO can restore the first-best outcome, free trade by setting $\bar{\tau} = 0$ and $Pr = 1$, that is, the optimization for V_C^W yields corner solutions.*

Proof. Since the WTO can access retaliatory tariff costlessly, the ideal mechanism can make every period retaliatory, while setting the lid as $\bar{\tau} = 0$. In this way, $V_C^W = \frac{u(0,0)}{1-\delta^C}$. □

Thus, a dilemma arises that either the WTO can do nothing, or it can do everything in the current setting. However, this is entirely due to the assumption that either the WTO does not have any verification ability as in the Proposition 5, or it effectively has perfect verification ability as in Proposition 6. However, before I proceed to

[103] By this assumption, the WTO effectively has an informational superiority besides the coordination function.

[104] A possible explanation might be that the country chooses to reveal the protection level in the punishment phase to the WTO, which requires a separately-specified mechanism. I discuss this issue further in Section 12.

the case where verification is costly, it is interesting to discuss about the case where the constraint is imposed such that any retaliatory tariff cannot be smaller than the cooperative one, that is, $\bar{\tau} \geq l$.

11.2. **Exact Bound.** Since I assume that the WTO can verify countries' actions during any retaliation phase, it is easier to start with analyzing the case where the WTO directly tells each country the exact protection level $\bar{\tau}$ that they should impose in a retaliation, instead of an upper bound. In particular, the following lemma implies that this assumption is useful for simplifying the analysis, by avoiding corner solution in $\bar{\tau}$ from above.

Lemma 5. *If $\bar{\tau} > h$, that is, the upper-bound of retaliation is above the static Nash level, then that bound does not bind. Otherwise, when the upper-bound is at most h, then it binds.*

Proof. h is dominant and the choice made in punishment phase does not lead to any triggering consequence. Thus, if $\bar{\tau}$ is only an upper bound, no countries implement it if $\bar{\tau} > h$.

□

Then, and throughout the rest of the paper, I assume that $\bar{\tau} \geq l$ and that the optimization in Section 11.1 always yields an interior solution for ω^D.[105] Thus it remains to be checked whether I can have interior/corner solutions for $\bar{\tau}$ and T^W, where a corner solution for $\bar{\tau}$ is either some lower bound for $\bar{\tau}$, or a prohibitive protection level, if there is any.[106] First, I consider the lower bound for $\bar{\tau}$. If $\bar{\tau}$ is set

[105]Or equivalently, I currently only consider the cases where the FOCs yield interior solution for ω^D. In some sense, this assumption rules out the possibility of $\bar{\tau} = 0$. Note that this assumption is satisfied in the numerical analysis later on.

[106]If there is no such level, a corner solution is $\bar{\tau} = \infty$.

at $\bar{\tau} = l$, then it is easy to check that $I^W(l) = 0$ only holds at $l = h$, implying a permanent Nash tariff war. Moreover, I have the following lemma.

Lemma 6. *The optimal restriction level $\bar{\tau}$ should be strictly greater than the equilibrium cooperative protection level.*

Proof. if $\bar{\tau}$ is set at $\bar{\tau} = l$, then $I^W(l) = \frac{\partial u(l,l)}{\partial \tau} = 0$ holds at $l = h$, implying a permanent Nash tariff war.

□

Then, I define the following two elasticities: the trigger control elasticity of sensitivity $\epsilon_1 \triangleq \frac{\partial^2 Pr(l)}{\partial \tau \partial \omega^D} \cdot \frac{\omega^D}{\frac{\partial Pr(l)}{\partial \tau}}$, and the trigger control elasticity of stability $\epsilon_2 \triangleq \frac{\partial [1-Pr(l)]}{\partial \omega^D} \cdot \frac{\omega^D}{[1-Pr(l)]}$,[107] which is useful in later analysis. In particular, it can be shown that:

$$\frac{dV_C^W}{dT^W} = \frac{\ln(\delta^C)\left(\delta^C\right)^{T^W}[1-Pr(l)][u(l,l)-u(\bar{\tau},\bar{\tau})]}{\left\{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]\right\}^2} \times \left[2Pr + \frac{\left(1-\delta^C\right)\cdot B\epsilon_2}{\left(\delta^C-(\delta^C)^{T^W}\right)B\epsilon_2 + A(\epsilon_1+\epsilon_2)}\right] \tag{11.6}$$

where both A and B are positive, and their expressions are given in the Appendix (a). Then, It can be shown the following proposition, which generalizes Park (2011)'s result regarding the choice of T^W. Since

$$\frac{\ln(\delta^C)\left(\delta^C\right)^{T^W}[1-Pr(l)][u(l,l)-u(\bar{\tau},\bar{\tau})]}{\left\{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]\right\}^2} < 0\,, \tag{11.7}$$

I obtain the following proposition.[108]

---

[107] Note that $\epsilon_1 > 0$, and $\epsilon_2 < 0$.

[108] All the conditions are evaluated at the optimal values of the control variables, unless otherwise noted.

Proposition 7. *Assume that the solution for ω^D is interior and that $l < \bar{\tau}$, then the optimal T^W is equal to 1, if and only if*

$$2Pr + \frac{\left(1-\delta^C\right)\cdot B\epsilon_2}{\left(\delta^C - \left(\delta^C\right)^{T^W}\right)B\epsilon_2 + A\left(\epsilon_1+\epsilon_2\right)} > 0. \tag{11.8}$$

Moreover, from Proposition 7, the following corollary holds:

Corollary 4. *If $\epsilon_1+\epsilon_2 \leqslant 0$, that is, the elasticity of sensitivity is less than the absolute value of the elasticity of stability, the optimal T^W is equal to 1.*

Proof. Since $Pr > 0$, $A > 0$, $B > 0$, $\epsilon_2 < 0$, the inequality $(\epsilon_1+\epsilon_2) \leqslant 0$ ensures that the condition in Proposition 7 is satisfied. □

This means that if the sensitivity is less responsive than the stability with respect to ω^D, then the shortest retaliation phase always results. Intuitively, it is because the social planner faces two trade-offs. Namely, it can either make the retaliation phase longer with less possibility of triggering retaliation, or make the retaliation phase shorter with greater possibility of triggering the retaliation. In addition, it faces the trade-off between sensitivity and stability. My results suggest that the duration of retaliation is more costly than the instability, after controlling for the sensitivity.

Corollary 5. *[Park (2011)] Assume that* $Pr(\tau) = \omega^D f(\tau)$ *for any* $\tau \geq 0$,[109] *then the optimal* T^W *is equal to 1, if and only if*

$$\frac{2Pr(l)}{1-\delta^C} + \frac{1}{\left(\delta^C - \left(\delta^C\right)^{T^W}\right) + \frac{A}{B} \cdot \frac{2Pr(l)-1}{Pr(l)}} > 0. \tag{11.9}$$

In particular, if $Pr(l) < \bar{Pr}$ *, where* $\left(1-\bar{Pr}\right)\left(1+\delta^C\right)+2\left(2\bar{Pr}-1\right)\left(1+\bar{Pr}\delta^C\right) = 0$,[110] *then the optimal* T^W *is equal to 1.*

Proof. see Appendix (b).

□

I have insofar assumed that the optimal ω^D is interior. Next I assume instead that the optimal $\bar{\tau}$ is interior. Then the following proposition obtains.

Proposition 8. *Assume that the optimal* $\bar{\tau}$ *is interior, then the optimal* T^W *is equal to 1, if and only if* $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial \tau}{\partial u(\bar{\tau},\bar{\tau})/\partial \tau^{\star}} < 1$. *In particular, if the optimal* $\bar{\tau} \leq h$, *then the optimal* T^W *is equal to 1.*

Proof. see Appendix (c).

□

Thus, it is hard to support both T^W and $\bar{\tau}$ as interior solutions simultaneously. Specifically, I have the following corollary:

---

[109] This condition yields $\frac{\partial^2 Pr(l)}{\partial \tau \partial \omega^D} > 0$, $\frac{\partial Pr(l)}{\partial \omega^D} > 0$ and $\frac{\partial^2 Pr(l)}{(\partial \omega^D)^2} = 0$, but the converse does not necessarily hold. In Park (2011), the assumption $\frac{\partial^2 Pr(\tau)}{(\partial \omega^D)^2} = 0$ guarantees that $Pr(\tau)$ is linear in ω^D, but it cannot ensure that $Pr(\cdot)$ goes through the origin in the $\omega^D - Pr$ space. And Park (2011) is actually using the assumption $Pr(\tau) = \omega^D f(\tau)$, instead of the one listed in his Proposition 3.

[110] $\bar{Pr}$ is also the upper bound in Park (2011)'s Proposition 3.

Corollary 6. *Whenever both the optimal T^W and the optimal $\bar{\tau}$ are interior, I must have $\bar{\tau} > h$, satisfying $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial \tau}{\partial u(\bar{\tau},\bar{\tau})/\partial \tau^{\star}} = 1$.*

In contrast with Park (2011)'s result where an interior T^W is possible, Corollary 6 implies that it is very unlikely that T^W is interior, whenever $\bar{\tau}$ is interior. Specifically, the duration of retaliation in my formulation is *weakly* shorter than the one in Park (2011) where $\bar{\tau} = h$. While the above-mentioned propositions imply that the WTO strictly prefers the additional instrument, $\bar{\tau}$, the following corollary establishes when, at $\bar{\tau} = h$, the WTO can help achieve strictly better outcome with the lid.

Corollary 7. *WTO can help achieve strictly better outcome with a lid in generic games, that is, whenever the optimal $\bar{\tau} \neq h$, or equivalently, $Pr(l) \neq \frac{\Delta_h}{2\Delta_h - 1}$, where $\Delta_h \equiv \frac{u(l,l)-u(l,h)}{u(l,h)-u(h,h)}$.*

Proof. see Appendix (c).

□

Thus, in general, WTO can do strictly better than the no-restriction case via this additional instrument. However, at this point, whether the WTO increases the lid above the static Nash protection level or put a lid lower than the static Nash level depends on different parametrizations. In this general setting, both of the two possibilities could exist. A quick remedy would be to conduct numerical analysis. Moreover, this exercise helps to understand the magnitude of the welfare improvement.

11.3. **A Numerical Example.** *Fixed discount factors*

I adopt the same parametrization as in Park (2011), except relaxing the constraint on the retaliatory protection level. Specifically, I assume the demand for good $i \in$

$\{1,2\}$ in country $j \in \{ ,\star\}$ is $D_i^j = A - Kp_i^j$, the supply for good i in country j is $X_i^j = \alpha_i^j + \kappa p_i^j$. Assume $\alpha_1 - \alpha_1^\star = \alpha_2^\star - \alpha_1 = 3$ with $\alpha_1 = \alpha_2^\star$, and $\kappa + K = 1$. Then, $u(\tau, \tau^\star)$ can be written as

$$u(\tau, \tau^\star) = \frac{1}{4}\left(3\tau - 3\tau^\star - \frac{3}{2}\tau^2 + \frac{1}{2}\tau^{\star^2}\right) \tag{11.10}$$

which yields a dominant $h = 1$.[111] In addition, I take the functional form of Pr as the following:

$$Pr(\tau) = \begin{cases} \omega^D\left[\frac{\tau^2}{2\chi} + \rho\right], & \text{for } \tau \le \sqrt{\omega^D\chi^{-1} - \chi\rho} \\ \omega^D\left[\frac{2\tau\cdot\sqrt{\omega^D\chi^{-1}-\chi\rho}}{\chi} - \frac{\tau^2}{2\chi} + 2\rho - \frac{1}{\omega^D}\right] & \text{for } \sqrt{\omega^D\chi^{-1} - \chi\rho} < \tau \le 2\sqrt{\omega^D\chi^{-1} - \chi\rho} \\ 1 & \text{for } \tau > 2\sqrt{\omega^D\chi^{-1} - \chi\rho} \end{cases} \tag{11.11}$$

where $\chi^{-1} \in (0,\infty)$, $\rho \in [0,\infty)$, and $\omega^D \in [0, \rho^{-1})$. ρ can be interpreted as the "noise in detection". For example, if $\rho = 0$, then $Pr(0) = 0$, implying that a zero protection level does not trigger the retaliation. As ρ becomes larger, there is more "noise" added into the model, making the detection even harder.

I let ρ varies from $80 \times (0.00005)$ to $130 \times (0.00005)$ and analyze the case when $\delta^C = 0.95$ (Figure 3.1) and $\delta^C = 0.5$ (Figure 3.2), respectively. In each figure, six panels correspond to, from top-left to bottom-right, (1) the Welfare, measured by V_C^W; (2) the protection level in the cooperative phase, l ; (3) the protection level in the punishment phase, $\bar{\tau}$; (4) the probability of triggering retaliation, $Pr(l)$; (5) the length of retaliation, T^W; (6) the trigger control variable, ω^D.

In Figure 3.1, the countries are patient with $\delta^C = 0.95$. First, note that the welfare gain from the lid is positive at every ρ. As shown in the figure, the optimal lid is interior with $\bar{\tau} < h$, which, according to Proposition 8, leads to a optimal $T^W = 1$, as

[111]Again, F's payoff can be written symmetrically.

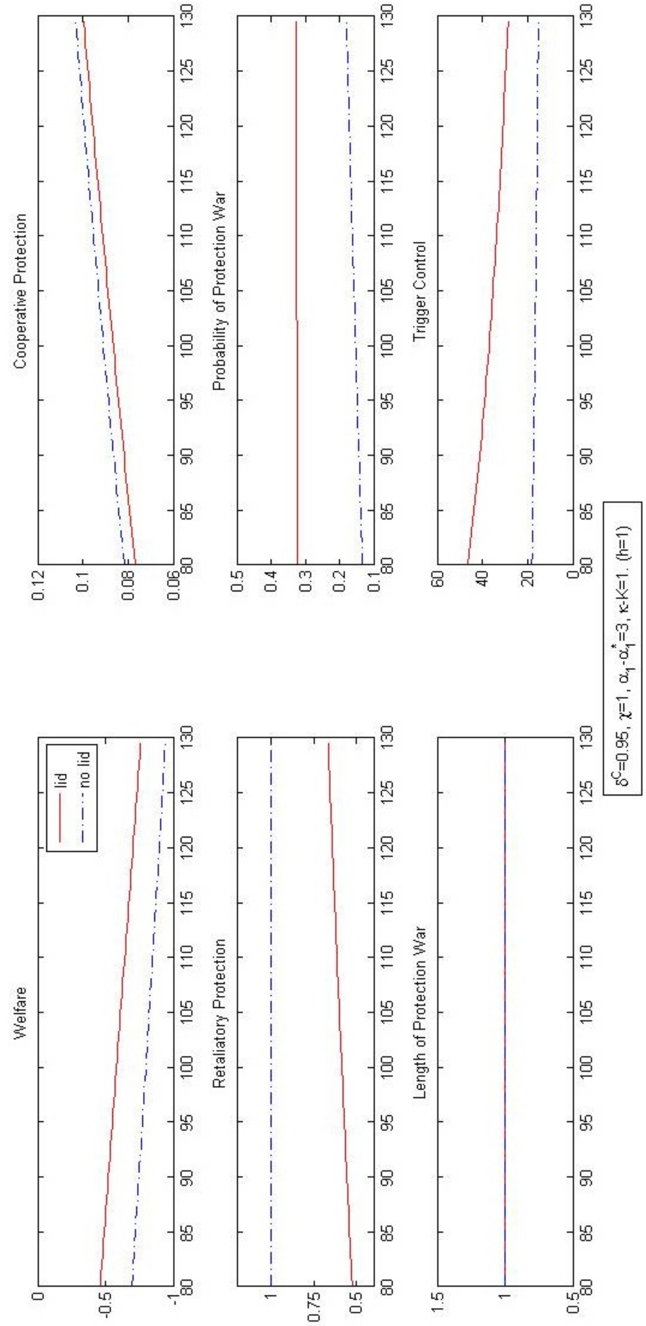

FIGURE 11.1. Patient Countries

depicted in the figure. A more surprising result is that even though $\bar{\tau}$ is below h, the

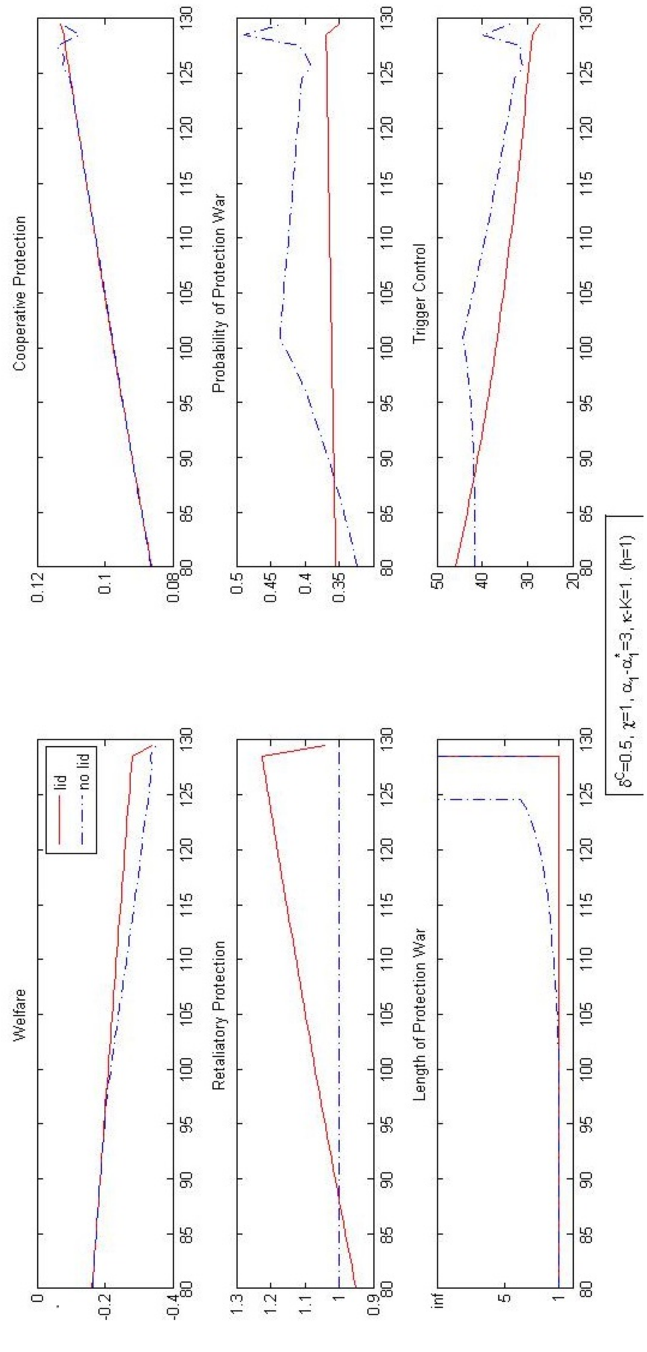

FIGURE 11.2. Impatient Countries

cooperative protection level is also slightly lower than the case in Park (2011). However, the triggering probability is much larger, which compensates the pain reduction from the retaliation phase, thus enhancing cooperation. In particular, the social planner increases the trigger control so that the sensitivity of detection increases, instead of using higher protection lid during the retaliation phase. However, the observation depends on specific values of parameters. In Figure 3.2, when countries are less patient with $\delta^C = 0.5$, I see that the optimal lid is above h for relatively large ρ. Thus, I conjecture that, other things equal, as δ^C increases, the social planner should rely more on the triggering device, rather than imposing a higher lid of retaliation, and as the noise of detection increases, a higher lid of retaliation is better. But a systematic analysis remains to be conducted so as to examine the actual relationship between different instruments to implement the retaliation. Also, in Figure 3.2, It can be seen that the welfare gain is larger when there are more noises in the detection, although this "noise effect" seems to disappear as countries become more patient.

Varying discount factors

Another important question involves whether free trade is attainable, i.e. $l = 0$, under the current framework for sufficiently patient countries.[112] The condition that at $l = 0$, $\frac{\partial u(l,l)}{\partial \tau} > 0$ implies that whether $I^W = 0$ holds depends on $\delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)][u(l,l) - u(l,\bar{\tau})]$. In particular, if $\frac{\partial Pr(l)}{\partial \tau} = 0$ at $l = 0$, which holds in my numerical analysis, then l is always larger than 0. Intuitively, for any given δ^C, countries always have an incentive to raise its l above zero, because there is no "marginal" punishment, given $\frac{\partial Pr(l)}{\partial \tau}|_{l=0} = 0$. To better understand the situation, I numerically consider this possibility next. Because it has been established that if the optimal $\bar{\tau} < h$, then $T^W = 1$, I fix $T^W = 1$ and $\rho = 100 \times (0.00005)$, in order to

[112]Park (2011) establishes that l cannot be zero for any given δ^C, but being silent about what if $\delta^C \to 1$.

consider what countries' welfare would look like when δ^C approaches zero. Then the optimization problem can be written as

$$\begin{aligned} &\max_{\omega^D,\bar{\tau}} \left\{ \frac{[1 - Pr(l)][u(l,l) - u(\bar{\tau},\bar{\tau})]}{1-\delta^C} + \frac{u(\bar{\tau},\bar{\tau})}{1-\delta^C} \right\} \\ &\text{s.t} \quad \frac{\partial u(l,l)}{\partial \tau} = \delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)][u(l,l) - u(l,\bar{\tau})] \end{aligned} \tag{11.12}$$

which is equivalent to the following problem:

$$\begin{aligned} &\max_{\omega^D,\bar{\tau}} \{[1 - Pr(l)][u(l,l) - u(\bar{\tau},\bar{\tau})] + u(\bar{\tau},\bar{\tau})\} \\ &\text{s.t.} \quad \frac{\partial u(l,l)}{\partial \tau} = \delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)][u(l,l) - u(l,\bar{\tau})] \end{aligned} \tag{11.13}$$

Thus, I am now able to consider the pattern of the economy around $\delta^C = 1$.[113] Specifically, the pattern is illustrated in Figure 3.3.

In Figure 3.3, although the cooperative protection is decreasing over δ^C, it never reaches zero as long as $\delta^C < 1$. Also note that $\bar{\tau} < h$ in my numerical example, and that for the no lid case I also checked that condition in Corollary 5 is satisfied for $\delta^C < 1$. Thus fixing $T^W = 1$ results in no bias in the analysis here.

However, an important question is not well addressed by the numerical analysis, that is, the reason why the WTO uses the lid. There are two possibilities: either a harsher retaliation enhancing more efficient cooperative protection, or a lesser retaliation reducing the welfare loss once the it is triggered. The current numerical example, at best, shows both of the two possibilities can exist depending on different sets of parameters.

Admittedly, in a sense, it is no surprise that the WTO can do better, because the WTO, by assumption, has an informational superiority in that it can effectively verify countries' concealed trade barriers in the punishment phase. This issue is explored

[113]The problem with the original optimization when δ^C approaches 1 is that the welfare tends to be unbounded.

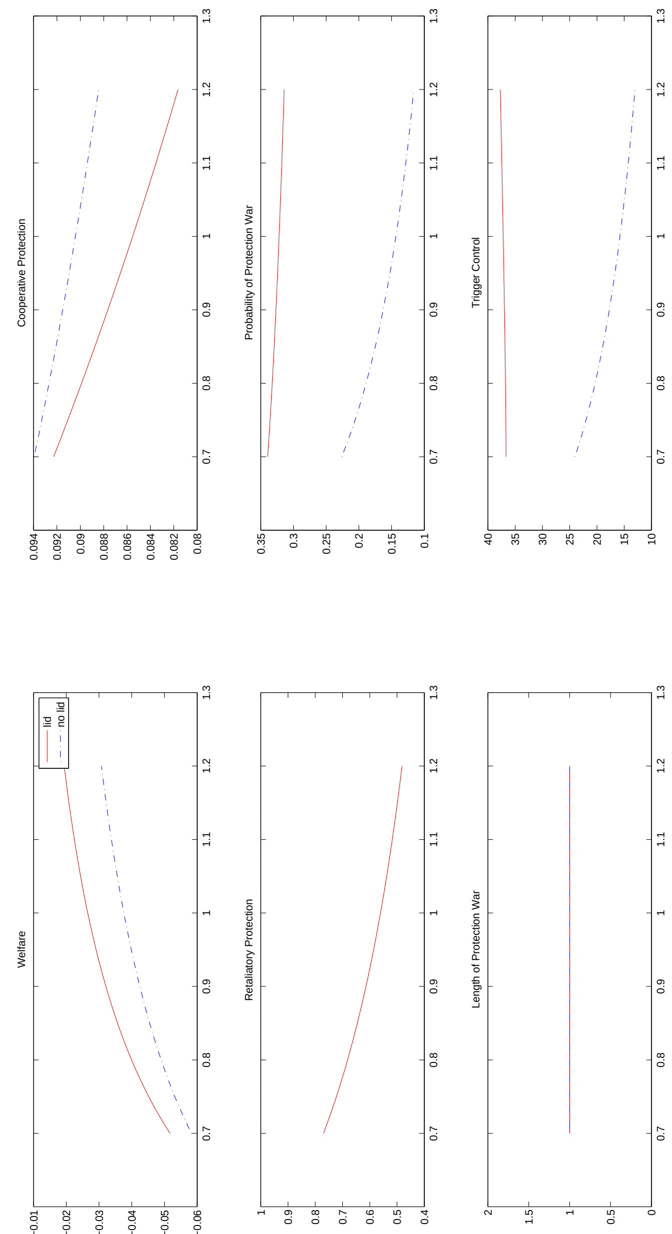

FIGURE 11.3. Sufficiently high δ^C

in Section 12 and Section 13. Specifically, in the next section, I introduce an option for member countries to pay a cost of establishing a panel to reveal, fully or partially, the concealed trade barriers to the WTO.

12. Verifiability of Retaliation: Cost of Establishing the Panel

12.1. **Call-upon Panel *versus* No Panel.** In Section 11, I simply assume that the WTO can directly verify the total protection levels during punishment phase, where I argue it is the Panel that makes such arrangements possible. However, the Panel is itself an endogenous choice by member countries. In practice, when the WTO receives the complaint, which can be treated as the signals received by the third party in the model, it establishes a panel to investigate the case. Here I assume such an establishment incurs a cost to member countries. In particular, in the two-country model, is it desirable for them to pay the "panel cost" such that the WTO has access to their total protection levels? If so, this is another possible mechanism that can explain why the WTO can/wants to put a lid on the retaliation level once the punishment is triggered.

In particular, I assume that every triggering is associated with a fixed cost of ϵ that is paid by both countries if only one country is being investigated and a fixed cost of 2ϵ if both countries are being investigated.[114] Thus, the new expected payoff function for H can be written as:

[114] This cost structure can preserve the condition$\delta^C - \delta^S = 2\left(\delta^C - \delta\right)$. If I assume a constant fixed cost ϵ (in both cases), my result is qualitatively similar.

(12.1)

$$V^P = (1-Pr)^2\left(u(l,l)+\delta^C V^P\right)$$
$$+Pr(1-Pr)\left(u(l,h)-\epsilon+\frac{\delta^C-\delta^{C^{T-1}}}{1-\delta^C}(u(h,h)-2\epsilon)+\delta^{C^{T-1}}\lambda(u(h,h)-2\epsilon)+\delta V^P\right)$$
$$+Pr(1-Pr)\left(u(h,l)-\epsilon+\frac{\delta^C-\delta^{C^{T-1}}}{1-\delta^C}(u(h,h)-2\epsilon)+\delta^{C^{T-1}}\lambda(u(h,h)-2\epsilon)+\delta V^P\right)$$
$$+Pr^2\left(u(h,h)-2\epsilon+\frac{\delta^C-\delta^{C^{T^S-1}}}{1-\delta^C}(u(h,h)-2\epsilon)+\delta^{C^{T^S-1}}\lambda^S(u(h,h)-2\epsilon)+\delta^S V^P\right)$$
$$\text{where } \delta \triangleq \lambda\delta^{C^T}+(1-\lambda)\delta^{C^{T-1}},\ \delta^S \triangleq \lambda^S\delta^{C^{T^S}}+\left(1-\lambda^S\right)\delta^{C^{T^S-1}}$$

Thus, the maximization problem facing the third party can be accordingly written as:

(12.2)

$$V^P\left(\delta^C,\epsilon\right) \triangleq \max_{\omega^D,T^W,\bar{\tau}}\left\{\frac{[1-Pr(l)][u(l,l)-u(\bar{\tau},\bar{\tau})]+2(1-Pr(l))\epsilon}{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]}+V_N^W\right\}$$

s.t. $I^W(l)=0$

$$\text{where}\quad I^W(l)\equiv\frac{\partial u(l,l)}{\partial\tau}-\delta^C\cdot\frac{\partial Pr(l)}{\partial\tau}\times [1-Pr(l)]\left\{u(l,l)-u(l,\bar{\tau})+\epsilon+\left(\delta^C-\left(\delta^C\right)^{T^W}\right)\left(V_C^W-V_N^W\right)\right\}$$

$$\text{and}\quad V_N^W\equiv\frac{u(\bar{\tau},\bar{\tau})-2\epsilon}{1-\delta^C}$$

$$\text{and}\quad V_C^W\equiv\frac{(1-Pr(l))[u(l,l)-u(\bar{\tau},\bar{\tau})]+2(1-Pr(l))\epsilon}{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]}+V_N^W$$

After comparing the above objective function with the one in Section 11, where there is no Panel at all with $\bar{\tau} = h$, it seems desirable for member countries to pay the cost resulted from the establishment of a Panel, as long as the fixed cost is not too large on the one hand. On the other hand, if the fixed cost is too large, establishing the Panel is not favorable. Formally, the following proposition is established, which provides sufficient conditions for the welfare improvement.[115]

Proposition 9. *Assume the optimal* $\bar{\tau}$ *is interior satisfying* $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial \tau}{\partial u(\bar{\tau},\bar{\tau})/\partial \tau^{\star}} < 1$*, then , other exogenous parameters equal, there exists some* $\epsilon_{np} > 0$ *such that if the fixed cost* $\epsilon \in [0, \epsilon_{np})$*, the WTO can help strictly improve the welfare via establishing the Panel once it is called upon in every retaliation phase. If* $\epsilon \in (\epsilon_{np}, \infty)$*, an option to establish the panel in the retaliation phase is not beneficial.*

Proof. See Appendix (d).

□

Although Proposition 9 establishes that countries invoke the Panel if the fixed cost is small, I am unable to analytically characterize the relationship between $\delta^C$and ϵ_{np} along the $V - V^P = 0$ locus at this point. In the numerical example, it is shown that $\frac{\partial \epsilon_{np}}{\partial \delta^C} > 0$.[116]

12.2. **Call-upon Panel *versus* All-time Panel.** I have insofar compared the the No Panel regime and Call-upon Panel regime. However, for a more satisfactory answer on the verifiability issue, it still remains to be discovered why there is no Panel in the cooperative phase within the current framework. Here I further explore the conditions under which countries choose to call upon the Panel only in the punishment phase rather than serve as a "policeman on the beat".

[115]This sufficient conditions largely reduces the complexity of calculation. And they are satisfied in the numerical example.
[116]From now on, V denotes the welfare without any lid, as in Park (2011).

Formally, if countries have an option to call upon the Panel at all time, the constraint obtained from countries' dynamic problem vanishes,[117] that is, the $I^W(l)=0$ condition in the above equations disappears. Thus their payoff can be simply written as

$$V^{ALL}\left(\delta^C,\epsilon\right)\equiv\max_{\omega^D,T^W,l,\bar{\tau}}\left\{\frac{[1-Pr(l)]\,[u(l,l)-u(\bar{\tau},\bar{\tau})]}{1-\delta^C+2Pr(l)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}+\frac{u(\bar{\tau},\bar{\tau})-2\epsilon}{1-\delta^C}\right\}=\frac{u(0,0)-2\epsilon}{1-\delta^C} \tag{12.3}$$

A comparison between $V^P\left(\delta^C,\epsilon\right)$ and $V^{ALL}\left(\delta^C,\epsilon\right)$ yields the conditions under which countries would choose to call upon a Panel.[118] Formally, I establish the following proposition:

Proposition 10. *(1) There exists some cutoff value $\hat{\epsilon}\left(\delta^C\right)>0$ such that for any $\epsilon>\hat{\epsilon}$, countries choose to call upon the Panel if and only if there is a dispute, i.e. in terms of the modeling terminology, only when punishment is triggered. (2) if δ^C is sufficiently large, or if the optimal $T^W=1$, then $\hat{\epsilon}\left(\delta^C\right)$ is decreasing over δ^C.*

Proof. See Appendix (e).

□

The above proposition implies a relationship between the fixed cost and the discount factor. Intuitively, on the one hand, when the fixed cost is too high, countries do not have the incentive to invoke the Panel, since the first purpose of a Panel is to reveal the actions during a punishment phase, which is served with priority over the second purpose as "policeman on the beat"; on the other hand, the higher is the

[117]Intuitively, it is because countries' actions can be perfectly observed now.

[118]Admittedly, I am implicitly assuming that countries only have access to two options: either they choose to pay the cost at all time or they pay only when any retaliation/punishment is triggered. A possible justification is that I am only focusing at the stationary equilibria.

discount factor, gain from cooperation is magnified more, which dominates the loss magnification during the punishment phase. Thus, countries have more incentive to cooperate, as a result, reducing the necessity to supervise them during the cooperation phase.

12.3. **All-time Panel *versus* No Panel.** A similar logic also applies in comparing All-time Panel and No Panel. When the fixed cost is high enough, which does not affect the welfare when there is no Panel, All-time Panel is not beneficial. Observing that this comparison is almost the same as the one in Section 12.2, I formally establish the following proposition likewise.

Proposition 11. *(1) There exists some cutoff value* $\tilde{\epsilon}\left(\delta^C\right) > 0$ *such that for any* $\epsilon < \tilde{\epsilon}$*, countries choose to establish a Panel at all time. (2) if* δ^C *is sufficiently large, or if the optimal* $T^W = 1$ *, then* $\tilde{\epsilon}\left(\delta^C\right)$ *is decreasing over* δ^C*.*

Proof. See Appendix (f).

□

Thus, Propositions 9-11 characterizes the relationship between different "panel regimes", in terms of $\left(\delta^C, \epsilon\right)$. Specifically, I have shown that countries choose among No Panel, Call-upon Panel and All-time Panel, depending on the fixed cost and their discount factors. Intuitively, when discount factors are sufficiently high, welfare in all three regimes increases, possibly at different rates. When the fixed cost increases, welfare in Call-upon Panel regime and All-time Panel regime decreases with No Panel regime unchanged. Given the analytical complexity to compare the rates at this point, I next conduct a numerical analysis in the next subsection.

12.4. **A Numerical Example.** I conduct a numerical analysis similar to the one in Section 11. Here I fix $\rho = 100 \times (0.00005)$, so as to examine the relationship between δ^C and ϵ. The cutoffs are presented in Figure 5.1. All locus correspond to $\Delta V = 0$, with different definitions of ΔV. As seen from the graph, both the $V^P - V^{ALL}$ locus and the $V - V^{ALL}$ locus displays the same pattern that δ^C and ϵ is negatively related, but the $V - V^P$ locus displays that δ^C and ϵ is positively related. And these three locus intersects at (0.836, 0.0232).

An interesting implication under this parametric specification is that as countries become more patient, the V^Pregime welfare increases fastest with V being the second and V^{ALL} being the last. Also, for a given discount factor, when the fixed cost is sufficient low, it is always beneficial for countries to invoke the Panel at all time. When the fixed cost is medium, countries switch to one of the two other regimes. When the fixed cost is high enough, countries definitely choose to be in the No Panel regime. Another interesting observation is that for a given fixed cost level, as long as countries are sufficiently patient, they end up in the Call-upon regime, due to the different effects δ^C has on the rate of welfare increase.

13. Verifiability of Retaliation: Imperfect Substitute between Explicit and Implicit Protection

In the previous analysis, explicit tariff and concealed trade barriers are perfect substitutes, where the explicit tariff plays no role, except that it serves as the coordination device/signal. Thus, as mentioned in Proposition 5, countries costlessly switch to concealed trade barriers if constrained in using the explicit tariff. However, as Copeland (1990) argues, it is not necessarily the case: generally speaking, concealed trade barriers are more costly to implement than the explicit tariff. Thus,

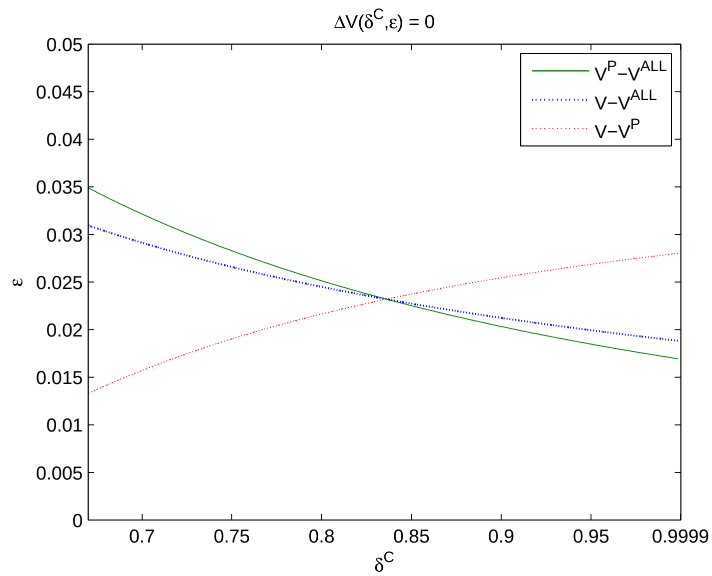

FIGURE 12.1. Cut-offs for Different Regimes

another possible reason why a lid may be implemented is that the two forms of barriers are imperfect substitutes to each other. Under this assumption, Proposition 5 may not necessarily hold.

Following Copeland (1990), I generalize the payoff function as

$$u(\tau, \tau^{\star}) \longrightarrow u(\tau, e, \tau^{\star}), \; u^{\star}(\tau^{\star}, \tau) \longrightarrow u(\tau^{\star}, e^{\star}, \tau)$$

which now depends on both the implicit and the explicit tariff. The protection package $a = (\tau, e)$ consists of the total protection level and the explicit tariff. A key assumption I make in the analysis is $u_{\tau e}(\tau, e, \tau^{\star}) > 0$ for $e < \tau$, which captures the idea that implicit tariff is more costly than the explicit one. Intuitively, it means that for any given total protection level τ' that the country wants to attain, the best choice is always setting $\tau = e = \tau'$. Since e plays no *actual* role in previous functions where the explicit tariff is not included in the payoff function, the original analysis regarding the welfare improvement still hold – as long as one modification is introduced into the framework. In particular, I need to use (l,l) as the signals for cooperation.[119]

Now I assume that the WTO is targeting at the explicit tariff, which is the pattern observed in practice. Now in the punishment phase, countries may not necessarily choose the static Nash h, since it might be that $u(\tau',\bar{e},\tau^\star) > u(h,\bar{e},\tau^\star)$ for some $\tau' < h$.[120] Thus, once the WTO puts a lid $\bar{e}$ on the explicit tariff, countries react to choose the best static choice in the punishment phase, specifically, countries act according to $\frac{\partial u(\bar{\tau},\bar{e},\bar{\tau})}{\partial \tau} = 0$. As a result, the new welfare optimization becomes:

$$\max_{\omega^D,T^W,\bar{e}} \left\{ \frac{[1-Pr(l)][u(l,l,l)-u(\bar{\tau},\bar{e},\bar{\tau})]}{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]} + V_N^W \right\} \tag{13.1}$$

$$\text{s.t.}\quad I(l) = 0$$

$$I^W(l) \equiv \begin{array}{c} \dfrac{\partial u(l,l,l)}{\partial \tau} - \delta^C \cdot \dfrac{\partial Pr(l)}{\partial \tau} \cdot [1-Pr(l)] \times \\ \left\{ u(l,l,l) - u(l,l,\bar{\tau}) + \left(\delta^C - \left(\delta^C\right)^{T^W}\right)\left(V_C^W - V_N^W\right)\right\} \end{array}$$

$$\text{where } V_N^W \equiv \frac{u(\bar{\tau},\bar{e},\bar{\tau})}{1-\delta^C},\ V_C^W \equiv \frac{(1-Pr(l))[u(l,l,l)-u(\bar{\tau},\bar{e},\bar{\tau})]}{1-\delta^C+2Pr(l)\left[\delta^C-(\delta^C)^{T^W}\right]} + V_N^W$$

$$\text{and } \frac{\partial u(\bar{\tau},\bar{e},\bar{\tau})}{\partial \tau} = 0$$

Note that by choosing $\bar{e}$, the lid is essentially targeting at the total protection level, $\bar{\tau}$ via the condition $\frac{\partial u(\bar{\tau},\bar{e},\bar{\tau})}{\partial \tau} = 0$. Again, while it might be difficult to pin down the optimal choice, I can still discuss the possibility of welfare improvement at around $(\tau,\bar{e}) = (h,h)$. Concretely, I have the following proposition:

[119] $\tau = l$ is the country's best choice as well.

[120] Note that $(h,h,\tau^\star)$ is still dominant.

Proposition 12. *When the implicit tariff and explicit tariff are imperfect substitutes in the sense given by the assumptions, by lowering the explicit tariff, the WTO can help countries strictly improve their welfare.*

Proof. See Appendix (g).

□

Thus, in this case, even though the WTO can not directly impose the restrictions on the concealed trade barriers, it can still help improve welfare due to the connection between implicit tariff and explicit tariff.

14. Concluding Remarks

A very important issue in the WTO framework involves the settlement of various disputes: on the one hand, it is desirable to reduce the welfare loss resulted from retaliation, if any; on the other hand, the threat needs to be high enough so as to enhance cooperation. This paper discusses about the possibility of welfare improvement when the WTO has more instruments, specifically when it can put restrictions on the retaliation levels chosen by countries once a punishment phase is triggered. It is argued that under verifiability of protection level during the punishment phase, in generic cases, the WTO can help achieve strictly better outcome than the case where such an instrument does not exist. I also argue that the Panel is the element that can make this possible. And the verifiability issue is discussed as well. In particular, my analysis not only provides a rationale for the WTO, but also lends support to the Panel. And, as long as the fixed cost is sufficiently low, it seems countries can always benefit from the Panel option: either a Call-upon regime or a All-time regime. Specifically, the numerical analysis seems to suggest that for any given fixed cost of a Panel, when countries are sufficiently patient, the Call-upon Panel regime is a best choice.

Meanwhile, for any given discount factor, when the fixed cost is sufficiently low, the economy ends up with free trade. Also, while modifying the WTO era framework so as to consider the various possibilities of verification devices, I have not yet considered these possibilities in an environment without the WTO. A shortcut answer would be that it is the WTO that brings these possibilities in an exogenous way. However, it is more interesting to augment the current settings so that they can be generated endogenously. Moreover, although the current setting is able to generate relatively low probability of retaliation, which seems to be consistent with the WTO practice, where retaliation is rare.[121] But it depends on how one maps the retaliations in the model into those in practice. For example, a Panel always corresponds to a retaliation phase in the current model, while it is not necessarily the case in the WTO's dispute settlement procedure. A well-defined mapping might also make it possible to calibrate or estimate the current model.

[121]See Bown and Ruta (2008) for the evidence.

15. Appendix

(a) Lagrangian for the optimization

Denote the Lagrangian function associated with the optimization as L, and the multiplier as λ. I establish that:

$$L = \left\{ \frac{\left[1 - Pr\left(l\right)\right]\left[u\left(l,l\right) - u\left(\bar{\tau},\bar{\tau}\right)\right]}{1 - \delta^C + 2Pr\left(l\right)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]} + \frac{u\left(\bar{\tau},\bar{\tau}\right)}{1 - \delta^C} \right\}$$

$$+ \lambda\{ \frac{\partial u\left(l,l\right)}{\partial \tau} - \delta^C \cdot \frac{\partial Pr\left(l\right)}{\partial \tau} \times$$

$$\left[1 - Pr\left(l\right)\right]\left\{ u\left(l,l\right) - u\left(l,\bar{\tau}\right) + \left[\delta^C - \left(\delta^C\right)^{T^W}\right]\left(V_C^W - V_N^W\right)\right\}\}$$

Thus, since I assume the optimal ω^D is always interior, the following first order condition should hold:

$$0 = \frac{\partial L}{\partial \omega^D} \equiv \frac{-\left[u\left(l,l\right) - u\left(\bar{\tau},\bar{\tau}\right)\right]\left(1 + \delta^C - 2\left(\delta^C\right)^{T^W}\right)\left[1 - Pr\left(l\right)\right]}{\left\{1 - \delta^C + 2Pr\left(l\right)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]\right\}^2} \cdot \frac{\partial Pr\left(l\right)}{\partial \omega^D}$$

$$+\lambda\left(-\delta^C\right)\cdot\{-\frac{\partial Pr\left(l\right)}{\partial \tau}\cdot\left[u\left(l,l\right) - u\left(l,\bar{\tau}\right) + \left(\delta^C - \left(\delta^C\right)^{T^W}\right)\left(V_C^W - V_N^W\right)\right] + \left[1 - Pr\left(l\right)\right]$$

$$\times\frac{\partial Pr\left(l\right)}{\partial \tau}\cdot\left(\delta^C - \left(\delta^C\right)^{T^W}\right)\cdot\frac{-\left[u\left(l,l\right) - u\left(\bar{\tau},\bar{\tau}\right)\right]\left(1 + \delta^C - 2\left(\delta^C\right)^{T^W}\right)\left[1 - Pr\left(l\right)\right]}{\left\{1 - \delta^C + 2Pr\left(l\right)\left[\delta^C - \left(\delta^C\right)^{T^W}\right]\right\}^2}\}\cdot\frac{\partial Pr\left(l\right)}{\partial \omega^D}$$

$$+\lambda\left(-\delta^C\right)\cdot\left[1 - Pr\left(l\right)\right]\cdot\left[u\left(l,l\right) - u\left(l,\bar{\tau}\right) + \left(\delta^C - \left(\delta^C\right)^{T^W}\right)\left(V_C^W - V_N^W\right)\right]\cdot\frac{\partial^2 Pr\left(l\right)}{\partial \tau \partial \omega^D}$$

Thus, letting $\epsilon_1 \equiv \frac{\partial^2 Pr(l)}{\partial \tau \partial \omega^D} \cdot \frac{\omega^D}{\frac{\partial Pr(l)}{\partial \tau}}$, $\epsilon_2 \equiv \frac{\partial[1-Pr(l)]}{\partial \omega^D} \cdot \frac{\omega^D}{[1-Pr(l)]}$, I have

$$\lambda\delta^C\frac{\partial Pr\left(l\right)}{\partial \tau}\left[1 - Pr\left(l\right)\right] = \frac{B\epsilon_2}{\left(\delta^C - \left(\delta^C\right)^{T^W}\right)B\epsilon_2 + A\left(\epsilon_1 + \epsilon_2\right)}$$

where

$$B \equiv \frac{\left(1+\delta^C-2\left(\delta^C\right)^{T^W}\right)\left[1-Pr\left(l\right)\right]}{\left\{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]\right\}^2}$$

$$A \equiv \Delta+\frac{\left(\delta^C-\left(\delta^C\right)^{T^W}\right)\left[1-Pr\left(l\right)\right]}{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}$$

$$\Delta \equiv \frac{u\left(l,l\right)-u\left(l,\bar{\tau}\right)}{u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)}$$

Moreover, I have the following partial derivatives:

$$\begin{aligned}\frac{\partial L}{\partial \bar{\tau}} =& \frac{-\left[1-Pr\left(l\right)\right]}{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}\cdot\frac{du\left(\bar{\tau},\bar{\tau}\right)}{d\bar{\tau}} \\ &+\frac{1}{1-\delta^C}\cdot\frac{du\left(\bar{\tau},\bar{\tau}\right)}{d\bar{\tau}}+\lambda\delta^C\frac{\partial Pr\left(l\right)}{\partial \tau}\left[1-Pr\left(l\right)\right] \\ &\times\left\{\frac{\partial u\left(l,\bar{\tau}\right)}{\partial \bar{\tau}}+\left(\delta^C-\left(\delta^C\right)^{T^W}\right)\cdot\frac{1-Pr\left(l\right)}{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}\cdot\frac{du\left(\bar{\tau},\bar{\tau}\right)}{d\bar{\tau}}\right\}\end{aligned}$$

$$\frac{\partial L}{\partial T^W}=\frac{\ln(\delta^C)\left(\delta^C\right)^{T^W}\left[1-Pr\left(l\right)\right]\left[u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)\right]}{\left\{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]\right\}^2}\left[2Pr+\lambda\delta^C\frac{\partial Pr\left(l\right)}{\partial \tau}\left[1-Pr\left(l\right)\right]\right]$$

(b) Proof for Corollary 5

If $Pr\left(\tau\right)=\omega^D f\left(\tau\right)$, then $\frac{\epsilon_1}{\epsilon_2}+1=\frac{2Pr(l)-1}{Pr(l)}$. Thus I obtain:

$$2Pr+\frac{\left(1-\delta^C\right)B\epsilon_2}{\left(\delta^C-\left(\delta^C\right)^{T^W}\right)B\epsilon_2+A\left(\epsilon_1+\epsilon_2\right)}=2Pr\left(l\right)+\frac{1-\delta^C}{\left(\delta^C-\left(\delta^C\right)^{T^W}\right)+\frac{A}{B}\cdot\frac{2Pr(l)-1}{Pr(l)}}.$$

In particular, if $Pr\left(l\right)\geq\frac{1}{2}$, then $\frac{dV_C^W}{dT^W}<0$. I then consider the case when $Pr\left(l\right)<\frac{1}{2}$. Specifically, I have the following logic:[122]

$$2Pr+\frac{1-\delta^C}{\left(\delta^C-\left(\delta^C\right)^{T^W}\right)+\frac{A}{B}\cdot\frac{2Pr(l)-1}{Pr(l)}}<0\Leftarrow$$

$$2Pr\delta^C+\frac{2\left(1+Pr\delta^C\right)\left(1-\delta^C+2Pr\delta^C\right)\left(2Pr-1\right)}{\left(1-\delta^C\right)\left(1-Pr\right)}+1-\delta^C<0$$

Note that $f\left(Pr\right)\triangleq 2Pr\delta^C+\frac{2\left(1+Pr\delta^C\right)\left(1-\delta^C+2Pr\delta^C\right)\left(2Pr-1\right)}{\left(1-\delta^C\right)\left(1-Pr\right)}+1-\delta^C$ is increasing over Pr. I only need to show that $f\left(\bar{Pr}\right)<0$. Since $\left(1-\bar{Pr}\right)\left(1+\delta^C\right)+2\left(2\bar{Pr}-1\right)\left(1+\bar{Pr}\delta^C\right)=0$, I have that

$$\begin{aligned}f\left(\bar{Pr}\right)=&2\bar{Pr}\delta^C-\frac{\left(1+\delta^C\right)\left(1-\delta^C+2\bar{Pr}\delta^C\right)}{1-\delta^C}+1-\delta^C\\&=\frac{-4\bar{Pr}\delta^{C^2}-2\delta^C+2\delta^{C^2}}{1-\delta^C}<0\end{aligned}$$

(c) Proof for Proposition 8 and Corollary 7

(Proposition 8) Given that $\bar{\tau}$ is interior, I must have $\frac{\partial L}{\partial\bar{\tau}}=0$. Based on the derivation in part (a), I obtain that:

$$\lambda\delta^C\frac{\partial Pr\left(l\right)}{\partial\tau}\left[1-Pr\left(l\right)\right]=$$

$$\left\{\frac{1-Pr\left(l\right)-\frac{1-\delta^C+2Pr(l)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}{1-\delta^C}}{\left\{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]\right\}\cdot\Theta+\left[1-Pr\left(l\right)\right]\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}\right\}$$

where $\Theta^{-1}\triangleq 1+\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}}$. Then, I can substitute it into $\frac{\partial L}{\partial T^W}$, and obtain:

$$\frac{\partial L}{\partial T^W}=\frac{\ln(\delta^C)\left(\delta^C\right)^{T^W}\left[1-Pr\left(l\right)\right]\left[u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)\right]Pr^2(l)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]}{\left\{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]\right\}^2\left(1-\delta^C\right)}\cdot(2\Theta-1)$$

[122] I impose that $\Delta=1$ and $T^W=\infty$, because the LHS is increasing in Δ and T^W when $Pr\left(l\right)<\frac{1}{2}$.

Thus, $\frac{dV_C^W}{dT^W}\left(=\frac{\partial L}{\partial T^W}\right)$ is negative as long as $\Theta > \frac{1}{2}$, or equivalently, $|\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}}| < 1$. At the same time, I have

$$\frac{\partial u\left(\bar{\tau},\bar{\tau}\right)}{\partial\tau}+\frac{\partial u\left(\bar{\tau},\bar{\tau}\right)}{\partial\tau^{\star}}<0$$

which implies that $|\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}}| < 1$ is equivalent to $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}} < 1$. In particular, if the optimal $\bar{\tau} = h$, then $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}} = 0 < 1$. Thus, the optimal T^W should be 1.

(Corollary 7) If the optimal $\bar{\tau} = h$, then $T^W = 1$. Substituting T^W into A and B and using $\frac{\partial L}{\partial\bar{\tau}} = 0$, I have that $Pr\left(l\right) = \frac{\Delta_h}{2\Delta_h - 1}$, with $\Delta_h \triangleq \frac{u(l,l)-u(l,h)}{u(l,h)-u(h,h)}$.

(d) Proof for Proposition 9

Corollary 7 establishes that the WTO generally can improve welfare strictly when $\epsilon = 0$, thus the welfare resulted from a fixed cost close to 0 is favorable. Formally, I can first show that $V^P\left(\delta^C,\epsilon\right)$ is decreasing over ϵ. Concretely, since ϵ is a fixed cost associated with $u\left(\cdot\right)$, the derivation in part (a) continues to hold except

$$\begin{aligned} A \triangleq & \Delta+\frac{\left(\delta^C-\left(\delta^C\right)^{T^W}\right)\left[1-Pr\left(l\right)\right]}{1-\delta^C+2Pr\left(l\right)\left[\delta^C-\left(\delta^C\right)^{T^W}\right]} \\ \Delta \triangleq & \frac{u\left(l,l\right)-u\left(l,\bar{\tau}\right)+\epsilon}{u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)+2\epsilon} \end{aligned}$$

Using Envelope Theorem, I have that:

$$\begin{aligned} \frac{\partial V^P}{\partial\epsilon} = & \frac{2\left(1-Pr\left(l^*\right)\right)}{1-\delta^C+2Pr\left(l^*\right)\left[\delta^C-\left(\delta^C\right)^{T^{W^*}}\right]}-\frac{2}{1-\delta^C}-\lambda\delta^C\frac{\partial Pr\left(l\right)}{\partial\tau}\left[1-Pr\left(l\right)\right] \\ & \times\left\{1+\frac{\left[\delta^C-\left(\delta^C\right)^{T^W}\right]\cdot 2\left(1-Pr\left(l^*\right)\right)}{1-\delta^C+2Pr\left(l^*\right)\left[\delta^C-\left(\delta^C\right)^{T^{W^*}}\right]}\right\} \end{aligned}$$

if the optimal $\bar{\tau}$ is interior satisfying $\frac{\partial u(\bar{\tau},\bar{\tau})/\partial\tau}{\partial u(\bar{\tau},\bar{\tau})/\partial\tau^{\star}}<1$, then $T^{W}=1$.[123] Thus,

$$\frac{\partial V^{P}}{\partial\epsilon}=\frac{2\left(1-Pr\left(l^{*}\right)\right)}{1-\delta^{C}}-\frac{2}{1-\delta^{C}}+\frac{Pr\left(l^{*}\right)}{\left(1-\delta^{C}\right)\Theta}$$

Thus, if $\Theta>0.5$, then $\frac{\partial V^{P}}{\partial\epsilon}<0$, implying $\frac{\partial\left(V^{P}-V\right)}{\partial\epsilon}<0$, where V is defined as the welfare without the lid, i.e. the one in Park (2011). Obviously, as ϵ becomes larger, V^{P} becomes negatively unbounded, while V is always bounded for any ϵ. Meanwhile, when $\epsilon=0$, I have already shown that $V^{P}>V$ in Corollary 7.

(e) Proof for Proposition 10

(1) Obviously, if $\epsilon=0$, I have

$$\begin{aligned}V^{P}\left(\delta^{C},0\right)&=\frac{\left[1-Pr\left(l^{*}\right)\right]\left[u\left(l^{*},l^{*}\right)-u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)\right]}{1-\delta^{C}+2Pr\left(l^{*}\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W*}}\right]}+\frac{u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)}{1-\delta^{C}}\\&<\frac{\left[u\left(l^{*},l^{*}\right)-u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)\right]}{1-\delta^{C}}+\frac{u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)}{1-\delta^{C}}\\&=\frac{u\left(l^{*},l^{*}\right)}{1-\delta^{C}}\\&\leq\frac{u\left(0,0\right)}{1-\delta^{C}}=V^{ALL}\left(\delta^{C},0\right)\end{aligned}$$

The last inequality holds since I have already shown $\frac{du(\bar{\tau},\bar{\tau})}{d\bar{\tau}}<0$ in Section 11. Intuitively, if the fixed cost of establishing a Panel is zero, countries simply do so at all time. And if this fixed cost becomes larger, it might not be welfare-improving to call upon the Panel at all time, as opposed to the case when Panel is only established at the punishment phase. Formally, I have

$$V^{P}\left(\delta^{C},\epsilon\right)-V^{ALL}\left(\delta^{C},\epsilon\right)=\frac{\left[1-Pr\left(l^{*}\right)\right]\left[u\left(l^{*},l^{*}\right)-u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)\right]+2\left(1-Pr\left(l^{*}\right)\right)\epsilon}{1-\delta^{C}+2Pr\left(l^{*}\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W*}}\right]}$$

$$-\frac{u\left(0,0\right)-u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)}{1-\delta^{C}}>\frac{2\left(1-Pr\left(l^{*}\right)\right)\epsilon}{1-\delta^{C}+2Pr\left(l^{*}\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W*}}\right]}-\frac{u\left(0,0\right)-u\left(h,h\right)}{1-\delta^{C}}$$

[123] Although ϵ is introduced here, the $\frac{\partial L}{\partial T^{W}}$ only differs by replacing $\left[u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)\right]$with $\left[u\left(l,l\right)-u\left(\bar{\tau},\bar{\tau}\right)+2\epsilon\right]$.

Observe that as ϵ becomes large enough, $V^{P}\left(\delta^{C},\epsilon\right)-V^{ALL}\left(\delta^{C},\epsilon\right)>0$. First, I can show that$V^{P}\left(\delta^{C},\epsilon\right)-V^{ALL}\left(\delta^{C},\epsilon\right)$ is increasing over ϵ. By the envelope theorem, I have that

$$\frac{\partial\left[V^{P}\left(\delta^{C},\epsilon\right)-V^{ALL}\left(\delta^{C},\epsilon\right)\right]}{\partial\epsilon}=\left(\partial V_{C}^{W}/\partial l\right)\cdot\frac{\partial l}{\partial\epsilon}+\frac{2\left(1-Pr\left(l\right)\right)}{1-\delta^{C}+2Pr\left(l\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W}}\right]}$$

$$=\left[-\frac{\partial V_{C}^{W}/\partial l}{\partial I^{W}(l)/\partial l}\right]\cdot\left[\partial I^{W}(l)/\partial\epsilon\right]+\frac{2\left(1-Pr\left(l\right)\right)}{1-\delta^{C}+2Pr\left(l\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W}}\right]}>0$$

$$(-)\qquad\qquad(-)\qquad\qquad(+)$$

(2) $\hat{\epsilon}\left(\delta^{C}\right)$ is given by $\triangle V\triangleq V^{P}\left(\delta^{C},\hat{\epsilon}\left(\delta^{C}\right)\right)-V^{ALL}\left(\delta^{C},\hat{\epsilon}\left(\delta^{C}\right)\right)=0$. Thus, by implicit function theorem, I have

$$\frac{\partial\hat{\epsilon}\left(\delta^{C}\right)}{\partial\delta^{C}}=-\frac{\partial\triangle V}{\partial\epsilon}/\frac{\partial\triangle V}{\partial\delta^{C}}$$

Given$\frac{\partial\triangle V}{\partial\epsilon}>0$, I only need to calculate the sign of $\frac{\partial\triangle V}{\partial\delta^{C}}$. First, if $T^{W}>1$, then

$$\frac{\partial\triangle V}{\partial\delta^{C}}=\left[-\frac{\partial V_{C}^{W}/\partial l}{\partial I^{W}(l)/\partial l}\right]\cdot\frac{\partial I^{W}\left(l\right)}{\partial\delta^{C}}+\Psi\cdot\left[\frac{1-2Pr\left(l\right)\left[1-T^{W}\left(\delta^{C}\right)^{T^{W}-1}\right]}{1-\delta^{C}+2Pr\left(l\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W}}\right]}-\frac{1}{1-\delta^{C}}\right]$$

where $\Psi\triangleq\dfrac{u\left(0,0\right)-u\left(\bar{\tau}^{*},\bar{\tau}^{*}\right)}{1-\delta^{C}}>0$

I already know $-\frac{\partial V_{C}^{W}/\partial l}{\partial I^{W}(l)/\partial l}<0$, and obviously, $\frac{\partial I^{W}(l)}{\partial\delta^{C}}<0$ since $\frac{1-\delta^{C}}{\delta^{C}-\left(\delta^{C}\right)^{T^{W}}}$ is decreasing. In addition,

$$\frac{1-2Pr\left(l\right)\left[1-T^{W}\left(\delta^{C}\right)^{T^{W}-1}\right]}{1-\delta^{C}+2Pr\left(l\right)\left[\delta^{C}-\left(\delta^{C}\right)^{T^{W}}\right]}/\left(\frac{1}{1-\delta^{C}}\right)=$$

$$\frac{1+2Pr\left(l\right)\left[T^{W}\left(\delta^{C}\right)^{T^{W}-1}-1\right]}{1+2Pr\left(l\right)\frac{\delta^{C}-\left(\delta^{C}\right)^{T^{W}}}{1-\delta^{C}}}\text{ as }\underset{\delta^{C}\to1}{\longrightarrow}\frac{1+2Pr\left(l\right)\left(T^{W}-1\right)}{1+2Pr\left(l\right)\left(T^{W}-1\right)}=1$$

Second, if $T^W = 1$, then

$$\frac{\partial \triangle V}{\partial \delta^C} = \left[-\frac{\partial V_C^W / \partial l}{\partial I^W (l) / \partial l} \right] \cdot \frac{\partial I^W (l)}{\partial \delta^C}$$

Thus, when δ^C is sufficiently high or $T^W = 1$, the sign of $\frac{\partial \triangle V}{\partial \delta^C}$ is determined by

$$\left[-\frac{\partial V_C^W / \partial l}{\partial I^W (l) / \partial l} \right] \cdot \frac{\partial I^W (l)}{\partial \delta^C} > 0$$

This implies $\frac{\partial \hat{\epsilon}\left(\delta^C\right)}{\partial \delta^C} < 0$.[124]

(f) Proof for Proposition 11

The proof is omitted here, as I just need to change $\bar{\tau}^*$ to h , and ϵ to 0 in $V^P$and I^W of the proof in part (e).

(g) Proof for Proposition 12

The derivative of the welfare function with respect to $\bar{e}$ can be written as:

$$\frac{dV_C^W}{d\bar{e}} = \frac{\partial V_C^W}{\partial \bar{\tau}} \cdot \frac{\partial \bar{\tau}}{\partial \bar{e}} + \frac{\partial V_C^W}{\partial l} \cdot \frac{\partial l}{\partial \bar{e}} + \frac{\partial V_C^W}{\partial \bar{e}}$$

In particular,

$$\frac{\partial V_C^W}{\partial \bar{\tau}} \cdot \frac{\partial \bar{\tau}}{\partial \bar{e}}|_{(\tau,\bar{e})=(h,h)} = \frac{\partial V_C^W}{\partial \bar{\tau}} \cdot \left[-\frac{\partial^2 u\left(\bar{\tau}, \bar{e}, \bar{\tau}\right)}{\partial \tau \partial e} / \frac{\partial^2 u\left(\bar{\tau}, \bar{e}, \bar{\tau}\right)}{\left(\partial \tau\right)^2} \right] < 0$$

$$\frac{\partial V_C^W}{\partial l} \cdot \frac{\partial l}{\partial \bar{e}}|_{(\tau,\bar{e})=(h,h)} = \left[-\frac{\partial V_C^W / \partial l}{\partial I^W (l) / \partial l} \right] \cdot \frac{\partial I^W (l)}{\partial \bar{e}} = 0$$

$$\frac{\partial V_C^W}{\partial \bar{e}}|_{(\tau,\bar{e})=(h,h)} = 0$$

Thus, $\frac{dV_C^W}{d\bar{e}}|_{(\tau,\bar{e})=(h,h)} < 0$.

(h) Verifiability of retaliation: a heuristic treatment

In Section 11, I assume that in the punishment phase, the WTO has access to the total retaliation level, which is more or less unnatural. In this appendix I explore

[124] Actually, the second term in $\frac{\partial \triangle V}{\partial \delta^C}$'s expression is negative, which implies that for not-so-large δ^C, the sign of $\frac{\partial \hat{\epsilon}\left(\delta^C\right)}{\partial \delta^C}$ is ambiguous.

the issue by endogenously generating countries' compliance with the WTO's command. An intuitive way would be to modify the basic framework such that there exist multiple Nash equilibria in the stage game, where I can induce various Nash-reversion options. However, while it is not impossible to conduct such analysis, this modification provides no help in comparing mine with Park (2011)'s results. Instead, given that in the punishment phase, private signals are used neither by countries in the PTS case, nor by the WTO in the TTS case, I argue that, by making use of that piece of information, the WTO may help to improve welfare, even without being able to access countries' retaliation levels. The idea is that the WTO can *punish* the countries, thus enforce the retaliation restriction, upon receiving their signals during the punishment phase. There might be several mechanisms that can be used, but I focus on augmenting the punishment phase of TTS in the following manner:

3'. If a *punishment* phase is initiated in period $t-1$, $\left(e_{t-1}, e^{\star}_{t-1}\right) \neq (0,0)$, then

- phase T : if $e_{t-1} \cdot e^{\star}_{t-1} = 0$, H plays $(\bar{\tau}, \bar{\tau})$ for the next $T-2$ periods and continues to play $(\bar{\tau}, \bar{\tau})$ for one more period with probability λ, repeat this phase with probability $Q_{\bar{\tau}}\left(\bar{\tau}, ..., \bar{\tau}\right)$.
- phase T^S: if $e_{t-1} \cdot e^{\star}_{t-1} \neq 0$, H plays $(\bar{\tau}, \bar{\tau})$ for the next T^S-2 periods and continues to play $(\bar{\tau}, \bar{\tau})$ for one more period with probability λ^S, repeat this phase with probability $Q^S_{\bar{\tau}}\left(\bar{\tau}, ..., \bar{\tau}\right)$.

where $Q^i_{\bar{\tau}}\left(\tau_1, ..., \tau_{T^i-1}\right)$ is the detecting/coordination device used by the WTO:[125]

$$Q^i_{\bar{\tau}}\left(\tau_1, ..., \tau_{T^i-1}\right) \equiv Pr\left(\left(\omega_1, ..., \omega_{T^i-1}\right) \in \bar{\Omega}^i_{\bar{\tau}} \middle| \left(\tau_1, ..., \tau_{T^i-1}\right), \left(\bar{\tau}, ..., \bar{\tau}\right)\right)$$

[125]For simplicity, I use $Q^i_{\bar{\tau}}$ and $Q^i_{\bar{\tau}}\left(\cdot\right)$ interchangeably when there is no confusion.

where $\bar{\Omega}^i_{\bar{\tau}}$ is the triggering set of repeating each phase, $i \in \{\ , S\}$.[126]

Then, due to modification of TTS, I need to re-calculate the V^W_C function. In particular, I have the following:

$$\begin{aligned}
V^W_C =& (1-Pr)^2\left(u(l,l)+\delta^C V^W_C\right)+Pr(1-Pr)\left(u(\bar{\tau},l)+\delta^C V_R\right)\\
&+Pr(1-Pr)\left(u(l,\bar{\tau})+\delta^C V_R\right)+Pr^2\left(u(\bar{\tau},\bar{\tau})+\delta^C V^S_R\right)\\
V_R =& \frac{1}{\delta^C}\left[\frac{\delta^C-\delta^{C^{T-1}}}{1-\delta^C}u(\bar{\tau},\bar{\tau})+\delta^{C^{T-1}}\lambda u(\bar{\tau},\bar{\tau})+\delta(1-Q_{\bar{\tau}})V^W_C+\delta Q_{\bar{\tau}}V_R\right]\\
=& \frac{1}{\delta^C}\left[\left(\delta^C-\delta\right)V^W_N+\delta(1-Q_{\bar{\tau}})V^W_C+\delta Q_{\bar{\tau}}V_R\right]\\
V^S_R =& \frac{1}{\delta^C}\left[\frac{\delta^C-\delta^{C^{T^S-1}}}{1-\delta^C}u(\bar{\tau},\bar{\tau})+\delta^{C^{T^S-1}}\lambda^S u(\bar{\tau},\bar{\tau})+\delta^S\left(1-Q^S_{\bar{\tau}}\right)V^W_C+\delta^S Q^S_{\bar{\tau}}V^S_R\right]\\
=& \frac{1}{\delta^C}\left[\left(\delta^C-\delta^S\right)V^W_N+\delta^S\left(1-Q^S_{\bar{\tau}}\right)V^W_C+\delta^S Q^S_{\bar{\tau}}V^S_R\right]\\
\text{where}\,&\delta \triangleq \lambda\delta^{C^T}+(1-\lambda)\delta^{C^{T-1}},\ \delta^S \triangleq \lambda^S\delta^{C^{T^S}}+\left(1-\lambda^S\right)\delta^{C^{T^S-1}}
\end{aligned}$$

I need to check that neither country has an incentive to cheat during the cooperation phase and the punishment phase.[127] For the cooperation phase I have that

$$Pr\left[u(\tau,\bar{\tau})+\delta^C V_R\right]+$$

$$(1-Pr)\left\{u(\tau,l)+\delta^C Pr(\tau)\left[u(l,\bar{\tau})+\delta^C V_R\right]+\delta^C[1-Pr(\tau)]\left[u(l,l)+\delta^C V^W_C\right]\right\}$$

$$Pr\left[u(\tau,\bar{\tau})+\delta^C V_R\right]+$$

$$(1-Pr)\left\{u(\tau,l)+\delta^C Pr(\tau)\left[u(\bar{\tau},\bar{\tau})+\delta^C V^S_R\right]+\delta^C[1-Pr(\tau)]\left[u(\bar{\tau},l)+\delta^C V_R\right]\right\}$$

I assume that $V^W_C - V_R = V_R - V^S_R$, which plays the same role here as the condition $\delta^C - \delta^S = 2\left(\delta^C - \delta\right)$ in Park (2011). Then, I obtain a similar constraint as $I^W(l)$ in

[126] The triggering set $\bar{\Omega}^i_{\bar{\tau}}$ may depend on the level of retaliation restrictions that the WTO wants to impose, which can help to capture that $\bar{\Omega}^i_{\tau}$=∅ for any $\tau \geq h$.

[127] In Park (2011), there is no incentive constraint involved in the punishment phase, because of the dominance of h.

Section 11:

$$\frac{\partial u(l,l)}{\partial \tau} = \delta^C \cdot \frac{\partial Pr(l)}{\partial \tau} \cdot [1 - Pr(l)] \left[u(l,l) - u(l,\bar{\tau}) + \delta^C \left(V_C^W - V_R\right)\right]$$

For the punishment phase, computation becomes messy, it requires that for any $j = 1, ..., T^i - 1$, the following conditions hold:

$$\left(\delta^C\right)^{j-1} \frac{\partial u(\bar{\tau},\bar{\tau})}{\partial \tau} = \frac{\partial Q^i_{\bar{\tau}}(\bar{\tau},...,\bar{\tau})}{\partial \tau_j}, \text{ for any } j < T^i - 1$$
$$\left(\delta^C\right)^{T^i-2} \lambda^i \frac{\partial u(\bar{\tau},\bar{\tau})}{\partial \tau} = \frac{\partial Q^i_{\bar{\tau}}(\bar{\tau},...,\bar{\tau})}{\partial \tau_{T^i-1}}. i \in \{ , S\}$$

Then, the social planner's problem is to choose $\left\{\omega^D, T^W, \bar{\tau}, Q^i_{\bar{\tau}}(\cdot)\right\}$ to maximize V_C^W:

$$\frac{2Pr(l)\delta^C\left(\delta^C - \delta\right)V_N^W + [1 - Pr(l)]\,u(l,l) + Pr(l)\,u(\bar{\tau},\bar{\tau})}{(\delta^C - \delta Q_{\bar{\tau}})(1 - \delta^C) + 2Pr(l)(\delta^C - \delta)}$$

subject to the constraint from the cooperation phase, the constraints from the punishment phase, and that $V_C^W - V_R = V_R - V_R^S$. Intuitively, at around $\bar{\tau} = h$, the punishment phase constraints is *marginally* satisfied,[128] the WTO again faces the choice between lowering$\bar{\tau}$ and lowering l. I expect that lowering $\bar{\tau}$ induces a first-order effect than the second-order effect from increasing l as in Section 11, that is, the WTO can still do strictly better in this more natural case. Formally, I can compute the sign of $\frac{dV_C^W}{d\bar{\tau}}|_{\bar{\tau}=h}$:

$$\frac{dV_C^W}{d\bar{\tau}}|_{\bar{\tau}=h} = \left[-\frac{\partial V_C^W/\partial l}{\partial I^W(l)/\partial l}\right] \cdot \frac{\partial I^W(l)}{\partial \bar{\tau}}|_{\bar{\tau}=h} + \frac{\partial V_C^W}{\partial \bar{\tau}}|_{\bar{\tau}=h}$$

Note that if $\bar{\tau} \geq h$, then the above problem exactly reduces to the case discussed in Park (2011). While still working on that issue, at this point, unfortunately I am unable to fully solve the problem and see how it works, in particular, I have not touched on the Q functions which is very important with regard to retaliation

[128] This depends on the choice/structure of $Q^i_{\bar{\tau}}(\cdot)$: $Q^i_{\bar{\tau}}(\cdot) = 0$ for any $\bar{\tau} \geq h$.

compliance. But the key idea is to make use of the signals from punishment phase to generate punishment on violating the retaliation restrictions via concealed protection level.

References

Admati, A. and Perry, M., 1987, "Strategic Delay in Bargaining", *Review of Economic Studies* 54(3), 345-364.

Bac, M. and Raff, H., 1997, "A Theory of Trade Concessions", *Journal of International Economics* 42(3-4), 483-504.

Bagwell, K., Mavroidis, P. and Staiger, R., 2007, "Auctioning Countermeasures in the WTO", *Journal of International Economics* 73(2), 309-332.

Bagwell, K. and Staiger, R., 1990, "A Theory of Managed Trade", *American Economic Review* 80(4), 779-795.

Bagwell, K. and Staiger, R., 1999, "An Economic Theory of GATT", *American Economic Review* 89(1), 215-248.

Bagwell, K. and Staiger, R., 2001, "Reciprocity, Non-discrimination and Preferential Agreements in the Multilateral Trading System", *European Journal of Political Economy* 17(2), 281-325.

Bagwell, K. and Staiger, R., 2002, *The Economics of The World Trading System*, MIT Press.

Bagwell, K. and Staiger, R., 2005, "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT/WTO", *Journal of International Economics* 67(2), 268-294.

Bagwell, K. and Staiger, R., forthcoming, "Multilateral Trade Bargaining and Dominant Strategies", *International Economic Review*.

Bagwell, K., Staiger, R. and Yurukoglu, A., 2015, "Multilateral Trade Bargaining: A First Look at the GATT Bargaining Records", NBER Working Paper No. 21488.

Bagwell, K., Staiger, R. and Yurukoglu, A., 2017, "Multilateral Trade Bargaining: A First Look at the GATT Bargaining Records", Working paper.

Binmore, K., Rubinstein, A. and Wolinsky, A., 1986, "The Nash Bargaining Solution in Economic Modelling", *RAND Journal of Economics* 17(2), 176-188.

Bown, C. and Ruta, M., 2008, "The Economics of Permissible WTO Retaliation", WTO Staff Working Paper ERSD-2008-04.

Broda, C., Limao, N. and Weinstein, D., 2008, "Optimal Tariffs and Market Power: The Evidence", *American Economic Review* 98(5), 2032-2065.

Cameron, A., Gelbach, J. and Miller, D., 2011, "Robust Inference With Multiway Clustering", *Journal of Business & Economic Statistics* 29(2), 238-249.

Cameron, A., and Miller, D., 2015, "A Practitioner's Guide to Cluster-Robust Inference", *Journal of Human Resources* 50(2), 317-372.

Cherry, J. and Smith, L., 2012, "Unattainable Payoffs for Repeated Games of Private Monitoring", mimeo.

Compte, O., 1998, "Communication in Repeated Games with Imperfect Private Monitoring", *Econometrica* 66(3), 597-626.

Copeland, B.,1990, "Strategic Interaction among Nations: Negotiable and Non-Negotiable Trade Barriers", *Canadian Journal of Economics* 23(1), 84-108.

Cramton, P., 1992, "Strategic Delay in Bargaining with Two-Sided Uncertainty", *Review of Economic Studies* 59(1), 205-225.

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

Fudenberg, D., Levine, D. and Tirole, J., 1987, "Incomplete Information Bargaining with Outside Opportunities", *Quarterly Journal of Economics* 102(1), 37-50.

Gawande, K. and Krishna, P., 2004, "The Political Economy of Trade Policy: Empirical Approaches", in *Handbook of International Trade*, Vol. 1, Choi, K. and Harrigan J. (eds.), Blackwell Publishing Ltd.

Gawande, K., Krishna, P. and Olarreaga, M., 2015, "A Political-Economic Account of Global Tariffs", *Economics & Politics* 27(2), 204-233.

Goldberg, P. and Maggi, G., 1999, "Protection for Sale: An Empirical Investigation", *American Economic Review* 89(5), 1135-1155.

Gomes, A., 2005, "Multilateral Contracting with Externalities", *Econometrica* 73(4), 1329–1350.

Green E. and Porter, R., 1984, "Noncooperative Collusion under Imperfect Price Information", *Econometrica* 52(1), 87-100.

Greene, W., 2004, "Fixed Effects and Bias Due to the Incidental Parameters Problem in the Tobit Model", *Econometric Reviews* 23(2), 125-147.

Grossman, G. and Helpman, E., 1995, "The Politics of Free-Trade Agreements", *American Economic Review* 85(4), 667-690.

Gu, W. and Kuhn, P., 1998, "A Theory of Holdouts in Wage Bargaining", *American Economic Review* 88(3), 428-449.

Harstad, B., 2007, "Harmonization and Side Payments in Political Cooperation", *American Economic Review* 97(3), 871-889.

Hoda, A., 2001, *Tariff Negotiations and Renegotiations Under the GATT and the WTO: Procedures and Practices*, Cambridge University Press.

Jehiel, P. and Moldovanu, B., 1995a, "Negative Externalities May Cause Delay in Negotiation", *Econometrica* 63(6), 1321-1335.

Jehiel, P. and Moldovanu, B., 1995b, Cyclical Delay in Bargaining with Externalities", *Review of Economic Studies* 62(4), 619-637.

Jensen, R. and Thursby, M., 1990, "Tariffs with Private Information and Reputation", *Journal of International Economics* 29(1-2), 43-67.

Kandori, M., 2002, "Introduction to Repeated Games with Private Monitoring", *Journal of Economic Theory* 102(1), 1-15.

Kandori, M. and Matsushima, H., 1998, "Private Observation, Communication and Collusion", *Econometrica* 66(3), 627-652.

Kennan, J., 1985, "The Duration of Contract Strikes in U.S. Manufacturing", *Journal of Econometrics* 28(1), 5-28.

Kennan, J. and Wilson, R., 1993, "Bargaining with Private Information", *Journal of Economic Literature* 31(1), 45-104.

Ludema, R. and Mayda, A., 2009, "Do Countries Free Ride on MFN?", *Journal of International Economics* 77(2), 137-150.

Ludema, R. and Mayda, A., 2013, "Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries", *Quarterly Journal of Economics* 128(4), 1837-1893.

Maggi, G. and Staiger, R., 2011, "The Role of Dispute Settlement Procedures in International Trade Agreements", *Quarterly Journal of Economics* 126(1), 475-515.

Matsushima, H., 2004, "Repeated Games with Private Monitoring: Two Players", *Econometrica* 72(3), 823-852.

Moser, C. and Rose, A., 2012, "Why Do Trade Negotiations Take So Long?", *Journal of Economic Integration* 27(2), 280-290.

Neary, P., 2004, "Europe on the Road to Doha: Towards a New Global Trade Round?", *CESifo Economic Studies* 50(2), 319-332.

Ossa, R., 2011, "A 'New Trade' Theory of GATT/WTO Negotiations", *Journal of Political Economy* 119(1), 122-152.

Park, J., 2011, "Enforcing International Trade Agreements with Imperfect Private Monitoring", *Review of Economic Studies* 78(3), 1102–1134.

Riezman, R., 1991, "Dynamic Tariffs with Asymmetric Information", *Journal of International Economics* 30(3-4), 267–283.

Rubinstein, A., 1982, "Perfect Equilibrium in a Bargaining Model", *Econometrica* 50(1), 97-109.

Smith, L. and Stacchetti, E., 2005, "Aspirational Bargaining", Working paper.

Sugaya, T., 2015, "Folk Theorem in Repeated Games with Private Monitoring", Working paper.

Tracy, J., 1987, "An Empirical Test of an Asymmetric Information Model of Strikes", *Journal of Labor Economics* 5(2), 149-173.

Yamamoto, Y., 2007, "Efficiency Results in N Player Games With Imperfect Private Monitoring", *Journal of Economic Theory* 135(1), 382-413.