CAEPR Working Paper
#2014-004

Arbitration and Renegotiation in Trade Agreements

Mostafa Beshkar
Indiana University

November 30, 2014

This paper can be downloaded without charge from the Social Science Research Network electronic library at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2533183

The Center for Applied Economics and Policy Research resides in the Department of Economics at Indiana University Bloomington. CAEPR can be found on the Internet at: <http://www.indiana.edu/~caepr>. CAEPR can be reached via email at caepr@indiana.edu or via phone at 812-855-4050.

©2014 by Mostafa Beshkar. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Arbitration and Renegotiation in Trade Agreements^{*†}

Mostafa Beshkar[‡]
Indiana University

November 30, 2014

Abstract

What can parties to a trade agreement achieve by institutionalizing a rules-based dispute settlement procedure? What types of renegotiation rules are implementable in an international setting like the WTO? What role can third-party arbitration play in dispute settlement? I study these questions within a mechanism design framework. The model generates predictions regarding the pattern of pre-trial and post-trial settlement negotiations, non-compliance with the arbitrator's ruling, and retaliations under an optimal trade agreement. It is shown that an Arbitrated-Liability Regime, under which a defecting party is liable for damages only to the extent that an arbitrator specifies, could implement the optimal direct mechanism.

^{*}I am grateful to Helen Milner, Robert Staiger, and Joel Watson for their in-depth discussion of this paper. Also, for their useful comments, I am grateful to Kyle Bagwell, Eric Bond, Kristy Buzard, Andrew Daughety, Johannes Horner, David Hummels, Nuno Limao, Giovanni Maggi, Roger Myerson, Jee-Hyeong Park, Jennifer Reinganum, and the seminar participants at Conference on the Economics, Law and Politics of the GATT/WTO at Yale University; Seoul National U., Vanderbilt, Purdue, U. New Hampshire, Midwest International Economics Meeting; Econometric Society's North American Summer Meeting; and the Workshop on Game Theory in Trade and Development in the Game Theory Festival, Stony Brook U.

[†]The first draft of this paper was circulated under the title "Third-Party-Assisted Renegotiation of Trade Agreements." Permanent link: <http://goo.gl/Hizho>.

[‡]Assistant Professor of Economics, Indiana University. Email: mbeshkar@indiana.edu.

KEYWORDS: Arbitration, Liability Rule, Property Rule, Safeguards, WTO
Dispute Settlement. (JEL CLASSIFICATION: F13, K33.)

1 Introduction

A remarkable achievement of the GATT/WTO negotiations has been the institutionalization of procedures to resolve disputes in implementation of trade agreements. Over time, dispute settlement among trading partners has evolved from a mostly informal bargaining process, with little restriction on acceptable norms of negotiations, to a rules-based process, with elaborate principles and rules for resolving disputes. This evolution towards a rules-based system was culminated in the Dispute Settlement Understanding (DSU) of the WTO, which specifies rules of negotiations and includes provisions for third-party arbitration.

Both the WTO and its predecessor, GATT, include an escape clause or safeguard agreement, which allows signatories to withdraw or modify trade policy concessions in response to unforeseen developments that cause or threaten serious injury to domestic producers. In the Law and Economics literature, the liability rule and the property rule specify different remedies for breach of a legal right. In the context of the WTO, the tariff commitments may be interpreted as conferring a legal right to low tariffs for exporting countries. However, the escape clause makes these rights effectively contingent on the state of the world, which may open the door to disputes in the implementation of the trade agreement if these contingencies are not easily verifiable.

The objective of this paper is three-fold. The first objective is to provide a model of the WTO's legal structure to analyze the pattern of dispute settlement and arbitration in the WTO. To this end, I interpret the legal structure of the WTO as an Arbitrated-Liability Regime (**ALR**), under which a party that has deviated from its obligations is *liable* for the resulting damages only to the extent that is determined by the arbitrator.¹ This interpretation is consistent with contingent

¹Liability Rule is a term that is used in the Law and Economics literature to describe a legal principle under which a contracting party who wishes to *escape* from its obligations is allowed to do so, but it will be also liable for damages that the defection inflicts on the other contracting parties.

protection provisions such as the WTO agreements on safeguards and antidumping measures. The Agreement on Safeguards, for example, stipulates contingencies under which a member country is entitled to violating its tariff commitments to a certain degree without being liable for the damages that it causes to the exporting countries. A safeguard-imposing country will be held liable only if it applies protection beyond what is authorized by the WTO Dispute Settlement Body (DSB). The magnitude of liability, which usually translates to retaliatory measures from exporting countries, is also determined by DSB arbitration based on remedy principles that are specified in the agreement.

Under this model, noncompliance with DSB rulings is an equilibrium outcome. This finding resonates with the empirical facts about noncompliance with the WTO rulings. In practice, in about one out of every five cases for which the DSB has issued a final ruling, the defending governments allegedly failed to bring their actions in compliance with the DSB recommendations immediately (Wilson 2007).²

My second objective is to study the impact of informal renegotiations on the optimal design of a rules-based legal system. I distinguish between formal (or rules-based) and informal renegotiations. A formal renegotiation is conducted under the auspices of the WTO and, as such, follows negotiation rules that are specified by the WTO. An example of negotiation rules is the GATT/WTO's Principle of Reciprocity, which posits that retaliation against a deviating country should be proportionate to the gravity of the original deviation. An informal renegotiation, in contrast, is conducted outside the realm of the WTO and is influenced by the existing norms of negotiation and power relations between the parties, which may differ from the norms promoted by the agreement.

This principle is usually compared to the Property Rule, under which violation of the contract is allowed only by consent of all the contracting parties.

²Examples of non-compliance with the WTO rulings include the Hormones dispute between the European Communities (as defendant) and the United States and Canada (as complainants), in which the European Communities declined to comply with the DSB's ruling to lift a ban on importation of beef products from US and Canada (WTO c). Another example is provided by the Canada-Dairy dispute, in which case the DSB's ruling against Canada was followed by a long period of negotiation between disputing parties. After more than three years of negotiations, the parties achieved a mutually accepted solution that was different from the original ruling of the DSB (WTO a).

A novel result of the paper is that an optimal dispute settlement procedure must authorize more severe retaliations against breaching countries that possess more bargaining power in informal renegotiations. To understand this result note that the outcome of arbitration determines the bargaining position of the parties, which could be different from the status quo under the agreement. Therefore, if the expected sanctions against breach are not sufficiently large, an importing country may have a better bargaining position under breach than under the status quo. Finally note that a party with greater bargaining power will gain more through ex post renegotiations. Therefore, to prevent an inefficient breach, the agreement should specify a harsher punishment against more powerful defectors so as to put them in a sufficiently bad bargaining position.³

This result indicates that an optimal rules-based system limits, but does not eliminate, the influence of the relative bargaining power of the parties. This paper, therefore, calls for more nuanced considerations in designing a rules-based dispute settlement procedure. In particular, designers of a legal institution must recognize that successful implementation of a rules-based system depends critically on the informal relations between the parties. This theoretical insight contributes to the debate regarding the limits of international law (e.g., Goldsmith and Posner 2005), by identifying the constraints that informal relations impose on the set of feasible rules.⁴

My third objective is to analyze the ‘optimality’ of the Arbitrated-Liability Regime (**ALR**), which was discussed above. To this end, I will set a benchmark by taking a mechanism design approach that is not restricted by the existing rules in the WTO. I will characterize the direct-revelation mechanism that maximizes the expected joint-

³Lyon and Rasmusen (2004) also consider a contractual relationship in which a party could propose an inefficient action to improve his bargaining position. They study the impact of relative bargaining power in ex post renegotiations and find that a buyer-option contract could prevent such inefficient actions.

⁴The model of this paper may be also extended to the study of ‘*decentralized law*’ in which the role of the legal system is to “find law [from community standards], rather than make it” (Cooter 1996). Similarly, the literature on the ‘*legal transplant*’ in developing countries (Mattei 1994; Buscaglia and Ratliff 2000) also recognizes the importance of a bottom-up approach in adopting laws. For example, Buscaglia and Ratliff (2000) argue that “the laws generating voluntary compliance are those that are truly compatible with the ethical code prevailing in a society.”

welfare of the governments under the assumption that the parties could renegotiate the outcome of the mechanism *ex post*. The resulting mechanism is an abstract creation with limited resemblance to the real-world institutions. I will then compare the outcome of this mechanism with that of the **ALR**, and find conditions under which the **ALR**, which resembles the general structure of the WTO agreement, is optimal.

This paper provides a framework to evaluate two canonical legal principles for breach remedies, namely, the Property Rule and the Liability Rule. Under the Property Rule, both parties' consent is required to alter the status quo policies that are specified by the arbitrator. In other words, the Property Rule leaves renegotiations unrestricted and parties will follow their informal norms to choose an outcome. In contrast, under the liability rule a defending party (here the breaching importing country) may deviate unilaterally from the arbitrator's recommended outcome as long as the complaining party (here the affected exporting country) is compensated according to a prespecified schedule. Therefore, in contrast to the Property Rule, renegotiations are structured by a prespecified retaliation scheme under the Liability Rule.

It is shown that the Property Rule is not an optimal *escape* provision. Under this rule the exporting country's consent is necessary for escape, which implies that the exporting country would be over-compensated, or at least fully compensated, for the breach of the agreement. But since compensation takes an inefficient form in trade relationships, it is optimal for an agreement to induce the minimum level of compensation that is necessary to prevent inefficient breach. Moreover, it is shown that the magnitude of compensation or retaliation that is necessary to ensure incentive compatibility, does not fully compensate the loss of the affected exporting countries. Hence, the Property Rule is suboptimal as an escape provision since it induces too much retaliation in the equilibrium.

This paper also sheds light on the apparent bias in the DSB rulings. As discussed by Sykes (2003), Grossman and Sykes (2007), and Colares (2009), in a strong majority of the cases the DSB rules at least partly against the defending party. The model suggests that this ruling pattern may be part of the optimal design of the system.

In particular, even if the DSB’s assessment of the disputed measure is in favor of the defending party, it is optimal for the DSB to authorize only a small deviation from the agreement tariff.

In a parallel research, Maggi and Staiger (2012) have developed a model of the DSB that generates pre-arbitration settlement and post-arbitration renegotiations. In contrast to the current paper, Maggi and Staiger (2012) assume that parties have symmetric information about the state of the world although this information is not verifiable to the court. Moreover, while I study continuous policies, Maggi and Staiger (2012) focus on disputes about trade policies that are discrete (binary) in nature.⁵

Beshkar (2010b) and Maggi and Staiger (2011) are among recent papers that provide formal models of the DSB.⁶ These models investigate alternative roles that an international tribunal like the DSB can play. As in this paper, Beshkar (2010b) shows the benefit of introducing arbitration into the WTO agreement by modeling the WTO as a signaling device. However, Beshkar (2010b) does not characterize the optimal agreement and does not allow for pre-arbitration settlement bargaining or ex post renegotiations. Beshkar (2010a) takes a mechanism-design approach but does not include the possibility of arbitration or renegotiations. The focus of Maggi and Staiger (2011) is on the role of arbitration when writing a complete contract is costly. Third-party arbitration, in their model, could improve cooperation by filling the gap in the contract under contingencies that are not specified in the agreement.⁷

In Section 2, I introduce the basics of the model, including the political-economy framework of the paper and the role of the DSB. In Section 3, I introduce a formal model of the Arbitrated-Liability Regime (**ALR**) and discuss the relevance of this

⁵Maggi and Staiger (2015) also study optimal trade agreements at the presence of renegotiations for the case where no informative arbitrator exists. I will discuss the relationship between my results and that of Maggi and Staiger (2015, 2012) in several points through out the paper.

⁶Earlier models of the WTO dispute settlement process include Reinhardt (2001), Ludema (2001), Rosendorff (2005), and Klimenko, Ramey, and Watson (2007). Park (2011) also provides a model of DSB as a public signaling device in a repeated-game framework. For a survey of these papers see Beshkar (2010b).

⁷Earlier models of the DSB include Maggi (1999), Rosendorff and Milner (2001), Reinhardt (2001), and Rosendorff (2005).

regime to the actual WTO agreement. Models of renegotiations in the interim and ex post stages are introduced in Section 4. Then, in Section 5, I set a benchmark for the rest of the paper by characterizing the optimal direct mechanism given the possibility of renegotiations. In Section 6, I show that the **ALR** could implement the outcome of the optimal direct mechanism. In Section 7, I study further implications of the optimal mechanism, including the role of informal norms/bargaining power on the optimal design of the agreement, and the optimality of the Property Rule and the Reciprocity Principle. Section 8 contains concluding remarks and suggestions for future research.

2 Basic Setting

I work within a simple political-economy trade framework that is used frequently in the literature.⁸ This framework is based on competitive markets with linear demand and supply functions in which countries gain from trade due to different production costs. The trade policy instrument at the governments' disposal is import tariffs. The details of this framework is laid out in Appendix A.

To focus on the problem of dispute resolution between two trading partners, I assume that there are two countries, a potential Defendant, denoted by D , and a potential Complainant, denoted by C . I will use τ and r to denote the import tariffs of D and C , respectively. Each government maximizes a weighted sum of its producers' surplus (π), consumers' surplus (ψ), and tariff revenues (R), with a potentially higher weight on the surplus of its import-competing sector.⁹ Denoting the political weight on the welfare of the import-competing sector by $\theta \geq 1$,¹⁰ D 's welfare drawn from its importable sector, m , is given by

$$u_D(\tau; \theta) \equiv \psi_m(\tau) + \theta \pi_m(\tau) + R(\tau),$$

⁸See, for example, Bagwell and Staiger (2005) and Beshkar (2010b).

⁹The higher weight given to the welfare of a sector may be the result of political pressure, through lobbying for example, that a government faces.

¹⁰In referring to θ , I will use "states of the world", "political-economy pressure", and "political-economy parameter" interchangeably.

where, τ is the specific tariff imposed on imports. The government's welfare from its exportable sector, x , is given by

$$v_D(r) \equiv \psi_x(r) + \pi_x(r).$$

where, r is the C 's import tariff. The payoffs of C may be defined in a similar way.

A simple way to model a dispute in this framework is to assume that the political-economy parameter of the potential defendant, D , is subject to privately-observed shocks. Since θ is not publicly observable, an upward adjustment in import tariffs as a response to an alleged increase in θ , may cause disagreement between the parties.

To capture the uncertainty in the future political-economy preferences, I assume that D 's political-economy parameter, θ , is drawn from a binary set $\{l, h\}$, $h > l \geq 1$, such that $\theta = h$ with probability ρ and $\theta = l$ with probability $1 - \rho$.¹¹ I sometimes use D_θ to refer to D of type θ .

For simplicity, I assume that C 's political-economy parameter is constant and equal to l .¹² Thus, letting $t \equiv (\tau, r)$, I use $V_C(t) \equiv u_C(r; l) + v_C(\tau)$ and $V_D(t; \theta) \equiv u_D(\tau; \theta) + v_D(r)$ to denote the payoffs of C and D , respectively. Then the joint welfare of the governments may be written as

$$W(t; \theta) \equiv V_D(t; \theta) + V_C(t).$$

The non-cooperative tariff of D , denoted by $\tau^N(\theta)$, is one that maximizes $u_D(\tau; \theta)$. The cooperative tariff of the D , denoted by $\tau^E(\theta)$, is defined as the tariff that maximizes the joint payoffs of the governments from D 's import tariff, i.e., $u_D(\tau; \theta) + v_C(\tau)$. Due to the terms of trade externality, the non-cooperative tariff is greater

¹¹I will assume that $h - l$ is positive but sufficiently small such that $\tau^N(l) \geq \tau^E(h)$. This assumption simplifies the analysis by eliminating the possibility of tariff binding overhang under an optimal agreement. For models of tariff overhang see Amador and Bagwell (2013), Beshkar, Bond, and Rho (2014), and Beshkar and Bond (2012).

¹²In practice, trade agreements are over multiple sectors, each of which may have a different political economy parameter. Moreover, these political economy parameters could be correlated across sectors and countries. However, as an initial attempt at modeling trade agreements in the presence of arbitration and renegotiations, I abstract from such possibilities by focusing on a single sector with uncertain political economy conditions.

than the cooperative tariff, i.e., $\tau^N(\theta) > \tau^E(\theta)$, as long as these tariffs are non-prohibitive. Nash and efficient tariffs of C , i.e., r^N and r^E , may be defined in a similar fashion.¹³ The tariff pairs that maximize the joint welfare of the governments will be denoted by $t^E(l) \equiv (\tau^E(l), r^E)$ and $t^E(h) \equiv (\tau^E(h), r^E)$. Finally, T denotes the set of tariff pairs, and $P_l, P_h \subset T$ denote the set of Pareto efficient tariff pairs under low and high political pressures, respectively.

I assume that tariffs are ‘*public actions*’, which are externally enforceable. In Watson’s (2007) terminology, a public action is one taken by an external enforcement entity, while an individual action is one that could be taken only by one of the contracting parties. It is common in the mechanism design literature to assume that all verifiable actions are public actions. However, Watson (2007) shows that this assumption may exclude some value functions that are otherwise implementable. Watson’s results are obtained for an environment in which: *i*) the state of the world is determined by unverifiable actions of the parties, *ii*) the trade actions are inalienable and irreversible, and *iii*) parties have ex post renegotiation opportunities. While satisfying the third condition, the present paper studies a different environment in which the state of the world is determined by an external shock rather than by actions of either parties. Moreover, the verifiable actions in this paper are tariffs that are essentially reversible.¹⁴

As in Beshkar (2010b) and Park (2011), I assume that the DSB is an impartial entity that receives a noisy signal, denoted by θ_A , about the state of the world in the defending country (Date 3-1). I assume that this signal matches the true state of the world, i.e., $\theta_A = \theta$, with probability $\gamma > \frac{1}{2}$, namely:

$$\Pr(\theta_A = l | \theta = l) = \Pr(\theta_A = h | \theta = h) = \gamma.$$

Assuming that the DSB has an informational role is broadly consistent with its mandate to make “objective assessment of the facts” of the dispute case and to make “recommendations” to help the disputing parties to develop a mutually satisfactory

¹³ $r^N \equiv \arg \max_r u_C(r; l)$, and $r^E \equiv \arg \max_r u_C(r; l) + v_D(r)$.

¹⁴More discussion of Watson’s critique is provided in the conclusion.

solution(WTO b). Through objective assessment of the facts, the DSB can obtain a signal, albeit imperfect, about the underlying political-economy conditions in the defending country. The recommendation of the DSB for a settlement, therefore, reflects the information that the DSB has obtained through its objective assessment.¹⁵

3 Arbitrated-Liability Regime

Under a liability-rule system, a contracting party who wishes to *escape* from its obligations is allowed to do so but it will be liable for resulting damages. As is well-known in the law and economics literature, an escape clause could improve the efficiency of a contractual relationship by encouraging *efficient breach* of the agreement, such that a contract is not performed under contingencies where the cost of performance exceeds its social benefits.¹⁶

In a similar fashion, the WTO members have reserved the right to increase their tariffs above their tariff binding commitments under certain conditions. As stated in Article 2 of the Agreement on Safeguards, a member country may apply a safeguard measure to a product only if a surge in imports “*cause or threaten to cause serious injury to the domestic industry.*” In practice, a safeguard measure is adopted by a government if the interested industries successfully lobby the government to take such actions. Through this process, the lobby groups are supposed to produce evidence regarding the extremely adverse impact of trade liberalization on their industries, which then may be used by the government to justify an escape from the agreement. Given the political-economy nature of this process, the safeguard provision may be interpreted as a safety valve for governments to diffuse political-economy pressures.¹⁷

¹⁵It is worth noting that assuming an informational role for the DSB does not imply any informational advantage on behalf of the DSB over the disputing parties. The advantage of the DSB over the disputing parties is its ‘impartiality’, which makes its public announcements about its privately observed signal reliable.

¹⁶As suggested by Sykes (1991, p. 284) and discussed below, an interesting analogy may be drawn between the liability regime and the WTO’s system of remedies. I extend his view by noting that under the WTO, the extent of liability depends on the prevailing contingency, which may be determined through third-party arbitration.

¹⁷Viewing safeguards as a means of diffusing political-economy pressures is a standard assumption

According the WTO agreement, the safeguard-imposing country has to exercise restraint in choosing the level of protection. Moreover, the affected countries may be entitled to some form of compensation. The WTO's Agreement on Safeguards (WTO 1995) states that

“A Member shall apply safeguard measures only to the extent necessary.... The affected exporting Members shall be free ... to suspend ... the application of substantially equivalent concessions. [However,] the right of suspension ... shall not be exercised..., provided that the safeguard measure ... conforms to the provisions of this Agreement.”

The method of compensation that is envisioned in this statement, is suspension of concessions, namely, retaliations, by the affected countries. Moreover, according to this clause, retaliations must be limited to withdrawal of substantially equivalent concessions.

I will also assume that tariff retaliations are the only compensation mechanism that is available at the time of implementing the agreement.¹⁸ However, instead of trying to quantify the level of substantially equivalent concessions, I will solve for the optimal level of retaliations. As I discuss in Section 7, it turns out that the reciprocity principle entails a suboptimal remedy mechanism.¹⁹

If the parties cannot agree on the size of necessary protection or the magnitude of retaliations, the WTO's dispute settlement body may be called for arbitration. We may interpret this system as an **Arbitrated-Liability Regime (ALR)**, since a defector is liable only to the extent that is determined by the arbitrator.

In order to build a formal model of the Arbitrated-Liability Regime, let t_b denote the negotiated tariff pair, which is supposed to be implemented in the normal times. Moreover, let t_s denote the tariff pair that is supposed to be implemented under the

in the literature. For further discussion of this point see Sykes (2006), Baldwin and Robert-Nicoud (2007). Section 2 of Beshkar (2010b) provides a summary.

¹⁸The WTO agreement does not rule out other methods of compensation. However, given that side payments are hard to come-by in practice, trade-policy retaliations are a more-practical form of compensation in settling disputes. For a discussion about the use of tariff retaliations vs. financial compensation in trade agreements see Limão and Saggi (2008).

¹⁹The magnitude of optimal retaliation depends on the relative bargaining power of the parties.

contingency in which a safeguard measure is justifiable. The Arbitrated-Liability Regime may be formally defined as follows:

Definition 1 *An Arbitrated-Liability Regime, **ALR**, is a trade agreement that is characterized by:*

- 1- A **baseline tariff** pair, $t_b \equiv (\tau_b, r_b)$, which determines the maximum tariff that each country is allowed to choose when political-economy pressures are low;
- 2- A **safeguard tariff** pair, $t_s \equiv (\tau_s, r_s)$, such that $\tau_s > \tau_b$ and $r_s = r_b$. τ_s is the maximum tariff that the importing country (i.e., Defendant) is allowed to choose when political-economy pressures are high, without facing retaliations ($r_s = r_b$);
- 3- An **arbitrator** who, in case of a disagreement between parties, will authorize escape if and only if $\theta_A = h$. Equivalently, the arbitrator recommends the tariff pair $t_A(\theta_A)$, such that $t_A(l) = t_b$ and $t_A(h) = t_s$.
- 4- A **retaliation scheme**, $r_{\theta_A}(\tau)$, which determines the magnitude of acceptable retaliation by C . In other words, C cannot choose any tariff above $r_{\theta_A}(\tau)$.

The time line of the **ALR** is shown in Figure 1. After observing its type, D could either apply $\tau_d \leq \tau_b$ or invoke the *escape clause* at Date 2. The choice of τ_b or the invocation of the escape clause may be interpreted as a message from D that potentially reflects its private information. If the escape clause is invoked, the game proceeds to arbitration (Date 3). The arbitrator first draws a signal, θ_A , at Date 3-1 and then authorizes the escape if and only it receives a high signal, $\theta_A = h$. If escape is authorized, D 's tariff binding increases from τ_b to τ_s . This implies that C 's tariff is bound at its baseline, r_b , unless retaliations are authorized. In other words, at Date 3-2 the arbitrator recommends $t_A(\theta_A) : \{l, h\} \rightarrow \{t_b, t_s\}$, and specifies the retaliation menu, $r_{\theta_A}(\tau)$, such that

$$\begin{aligned} t_A(l) &= t_b, \quad t_A(h) = t_s, \\ \tau_d &> \tau_A(\theta_A) \iff r_{\theta_A}(\tau_d) > r_b, \\ r_l(\tau_b) &= r_h(\tau_s) = r_b. \end{aligned}$$

The last set of equalities indicate that the arbitrator authorizes retaliations only if

Figure 1: Timeline of the Arbitrated-Liability Regime (**ALR**).

D applies a tariff in excess of the recommended level, i.e. if $\tau_d > \tau_A \in \{t_b, t_s\}$. Therefore, the retaliation menus will satisfy $r_l(\tau_b) = r_h(\tau_s) = r_b$.

In order to examine the optimality of the Arbitrated-Liability Regime, I first take a more general approach by finding the optimal direct mechanism under the same environment (Section 5). In principle, the optimal direct mechanism could result in a higher welfare than **ALR**. Nevertheless, as I show in Section 6, the **ALR** could in fact implement the outcome of the optimal direct mechanism.

A central assumption of the paper is that the trading partners could renegotiate the terms of the agreement both at the interim stage (Date 1) and the ex post stage (Date 4). The outcome of any mechanism, therefore, depends on the details of the bargaining procedures for each of these stages, which are laid out in the next Section.

4 Renegotiations

I consider two renegotiation possibilities: interim renegotiations (Date 1), which takes place after the realization of private information and before sending messages, and ex post renegotiation (Date 4), which takes place after the outcome of the mechanism is

Figure 2: Pareto improvement possibilities for a given tariff pair, t_d , when $\theta = l, h$.

determined. Both of these renegotiations take place under imperfect information.²⁰ In this section, I will discuss the bargaining models that I use for each of these renegotiation possibilities.

These two renegotiation possibilities take place in fundamentally different settings. In the interim stage, negotiations takes place in lieu of arbitration's uncertain outcome. In the ex post stage, however, the arbitration outcome has been already determined and the objective of renegotiations is to exploit any Pareto improvement that might be possible given the arbitrator's recommended outcome.

4.1 Ex Post Renegotiations

Suppose that countries C and D are currently implementing an externally-enforceable tariff pair, t_d , as depicted in Figure (2). In this figure, P_l and P_h depict the set of Pareto efficient tariff pairs when $\theta = l$ and $\theta = h$, respectively. The disagreement tariff pair, t_d , could be thought of as the outcome of a mechanism, which is enforceable by assumption.

²⁰Other papers in the literature, including Maggi and Staiger (2012), Maggi and Staiger (2015), study renegotiations under perfect information and use the Nash Bargaining Solution to model bargaining.

Figure 3: Timeline of Interim and Ex Post Renegotiations.

I model ex post renegotiations as a two-stage alternating-offer bargaining under asymmetric information. Formally, Suppose that the bargaining is over setting tariffs for a duration of one unit of time.²¹ At the beginning of the ex post negotiations, C proposes a menu of tariff pairs, M . If D chooses a tariff pair from the proposed menu, $t_p \in M$, an agreement is reached immediately and t_p will be applied for the entire time interval. Otherwise, the disagreement tariff, t_d , will be applied for a duration of $\Delta < 1$. At time Δ , D will have the opportunity to propose a tariff pair, t'_p , which may be accepted or rejected by C .

If the second stage is reached, C will accept t'_p if and only if $V_C(t'_p) \geq V_C(t_d)$. Therefore, the best strategy for D of type θ in the second stage is to propose $t'_p = t_p^{\max}$ such that

$$t_p^{\max}(\theta) \in P_\theta, \quad V_C(t_p^{\max}) = V_C(t_d). \quad (1)$$

By backward induction, this implies that D will accept a proposed tariff t_p in the

²¹This assumption is consistent the WTO requirement that safeguards must be temporary.

first stage only if

$$\int_0^1 e^{-rz} V_D(t_p; \theta) dz \geq \int_0^\Delta e^{-rz} V_D(t_d; \theta) dz + \int_\Delta^1 e^{-rz} V_D(t_p^{\max}(\theta); \theta) dz, \quad (2)$$

where, r is D 's discount rate. The left-hand side of this condition is the discounted payoffs of D if an agreement is achieved without a delay. The right-hand side of condition (2) is the discounted payoffs of D if agreement is achieved only after a delay of Δ . The first expression on the right-hand side is the discounted payoffs of the *disagreement* period, i.e., until time Δ . The second expression is the discounted payoffs from time Δ to 1, during which time $t_p^{\max}$ is applied. Defining $\lambda \equiv \int_\Delta^1 e^{-rz} dz / \int_0^1 e^{-rz} dz$, condition (2) can be simplified to:

$$V_D(t_p; \theta) \geq (1 - \lambda) V_D(t_d; \theta) + \lambda V_D(t_p^{\max}(\theta); \theta). \quad (3)$$

The parameter λ , which is decreasing in Δ and increasing in r , captures the bargaining power of C .

It turns out that the outcome of bargaining under asymmetric information depends critically on the initial point of bargaining, t_d . In particular, depending on t_d and C 's prior about the true state of the world, the ex post bargaining may or may not reach a Pareto efficient outcome. For a subset of tariff pairs, however, asymmetry of information does not affect the outcome of ex post renegotiations.

To elaborate, let $\Upsilon_\lambda^\theta : T \longrightarrow T$ be a mapping such that

$$\Upsilon_\lambda^\theta(t_d) \in P_\theta, \quad (4)$$

$$V_D(\Upsilon_\lambda^\theta(t_d); \theta) = (1 - \lambda) V_D(t_d; \theta) + \lambda V_D(t_p^{\max}(\theta); \theta). \quad (5)$$

In other words, $\Upsilon_\lambda^\theta(t_d)$ is defined as a Pareto efficient tariff pair given θ , which makes D_θ indifferent between implementing $\Upsilon_\lambda^\theta(t_d)$ immediately and delaying an agreement

for Δ units of time. Moreover, let $T^* \subset T$ be the set of all tariff pairs, t_d , such that

$$\Upsilon_\lambda^h(t_d) \stackrel{D_l}{\succeq} \Upsilon_\lambda^l(t_d), \quad (6)$$

$$\Upsilon_\lambda^l(t_d) \stackrel{D_h}{\succeq} \Upsilon_\lambda^h(t_d), \quad (7)$$

where $\stackrel{D_\theta}{\succeq}$ indicates the preference relation of D_θ . Then,

Proposition 1 *In the ex post renegotiations, C proposes the menu $M \equiv \{\Upsilon_\lambda^l(t_d), \Upsilon_\lambda^h(t_d)\}$, and D_θ accepts $\Upsilon_\lambda^\theta(t_d)$ if and only if $t_d \in T^*$.*

Proof. First note that by definition, D_θ is indifferent between accepting $\Upsilon_\lambda^\theta(t_d)$ and rejecting a deal at Date 4-2. Moreover, conditions 6 and 7 ensure that D_θ (weakly) prefers $\Upsilon_\lambda^\theta(t_d)$ to $\Upsilon_\lambda^{\theta'}(t_d)$. Therefore, if the menu $M \equiv \{\Upsilon_\lambda^l(t_d), \Upsilon_\lambda^h(t_d)\}$ is proposed by C , a low-type D will accept $\Upsilon_\lambda^l(t_d)$ and a high-type D will accept $\Upsilon_\lambda^h(t_d)$, immediately.

Now consider C 's decision to propose a menu at Date 4-1. C prefers both $\Upsilon_\lambda^l(t_d)$ and $\Upsilon_\lambda^h(t_d)$ to no deal at Date 4-2. Moreover, Given that $\Upsilon_\lambda^l(t_d) \in P_l$ is Pareto efficient and it makes D_l just indifferent about accepting the offer, there is no alternative tariff pair that improves C 's welfare and induces D_l to accept the proposal immediately. The same argument applies to $\Upsilon_\lambda^h(t_d)$. Thus, C will propose $M \equiv \{\Upsilon_\lambda^l(t_d), \Upsilon_\lambda^h(t_d)\}$ if conditions 6 and 7 are satisfied.

To show that conditions 6 and 7 are also necessary for this outcome, note that if either of these conditions is violated, then the two types will pool by choosing the same tariff from the proposed menu $M \equiv \{\Upsilon_\lambda^l(t_d), \Upsilon_\lambda^h(t_d)\}$. QED. ■

4.2 Interim Renegotiations

I consider the possibility of renegotiations at the interim stage, defined as the time between the (private) realization of the state of world and executing the status quo mechanism (Holmström and Myerson 1983). Similar to Beaudry and Poitevin (1995), I assume that one party has the opportunity to offer an alternative mechanism at this stage. The interim renegotiation problem, therefore, could be thought of as a new

mechanism-design problem where each party knows its type and the outside option of the parties is given by the status quo mechanism. This mechanism is restricted by the same constraints as in the main mechanism-design problem, namely, incentive compatibility and renegotiation-proof constraints.

Formally, C proposes an alternative mechanism at date 1-1 (Figure 3). At Date 1-2, D may accept or reject this proposal. If D rejects the proposal, the game proceeds to Date 2 and the status quo mechanism will be executed. If D accepts the proposal, the alternative mechanism will be executed. In order to study the outcome of interim renegotiations, we first need to calculate each party's expected welfare from playing out the (status quo) mechanism. In the subsequent sections, the interim renegotiations will be studied as part of the optimal mechanism design problem.

5 The Optimal Direct-Revelation Mechanism

In this section, I study the design of trade agreements as a renegotiation-proof direct mechanism.²² This mechanism will set a useful benchmark for the subsequent sections in which I study the common legal institutions and rules such as the liability rule and the property rule systems.

The complete timeline of the game induced by a direct mechanism is depicted in Figure 4. In the messaging stage (Date 2), D directly reports its type. In the arbitration stage, instead of assuming a structure such as the **ALR**, I assume that the DSB could determine an enforceable outcome as a function of D 's report, θ_D , and its own observed signal, θ_A . I, therefore, use $t_{\theta_D\theta_A} \equiv (\tau_{\theta_D\theta_A}, r_{\theta_D\theta_A})$ to denote a mechanism or decision rule.

²²Brennan and Watson (2013) show that the possibility of renegotiation amounts to a constraint on the problem. They formalize a 'Renegotiation-Proofness Principle' and find conditions under which this principle holds. This principle states that any payoff vector that is implementable with renegotiation can also be implemented by a mechanism that is renegotiation proof. However, in general the renegotiation-proofness requirement may preclude some payoff vectors that are implementable. I follow a strand of the literature (e.g., Dewatripont 1989) that focuses on renegotiation-proof mechanisms.

Figure 4: Complete timeline of the game induced by DM.

I assume that the mechanism designer's objective is to maximize the expected joint welfare of the governments defined as $W(t; \theta) \equiv V_D(t; \theta) + V_C(t)$.²³ If a decision rule, $t_{\theta_D \theta_A}$, is implemented, the expected joint welfare of the parties will be given by

$$EW(t_{\theta_D \theta_A}) \equiv \rho [\gamma W(t_{hh}; h) + (1 - \gamma) W(t_{hl}; h)] + (1 - \rho) [\gamma W(t_{lh}; l) + (1 - \gamma) W(t_{ll}; l)]. \quad (8)$$

Given the focus of the paper on renegotiation-proof mechanisms, we can restrict attention to a subset of tariff pairs that could be part of a renegotiation-proof mechanism. In particular:

Lemma 1 *Any $t_d \notin P_l \cup P_h$ will be renegotiated in the ex post stage regardless of the true state of the world. Moreover, $P_l \cup P_h \subset T^*$.*

²³This assumption is plausible if countries are ex ante symmetric or if governments can transfer side payments at the time of crafting a trade agreement. It is also straightforward to use a weighted joint-welfare. However, since the structure of the optimal agreement is determined by the incentive compatibility constraints, a weighted joint welfare function does not change our qualitative results.

Proof. Suppose that, on the contrary, $t_d \notin P_l \cup P_h$ is not renegotiated at Date 4-1. Then at date 4-2, D of type θ proposes $t'_p = t_p^{\max}(\theta) \in P_\theta$ and C accepts the offer. Therefore, $t_d \notin P_l \cup P_h$ will be renegotiated in the ex post stage for any true state of the world. It is straightforward to show that any tariff pair in $P_l \cup P_h$ satisfies conditions 6 and 7. Hence, $P_l \cup P_h \subset T^*$. ■

This lemma implies that a truthful mechanism is impervious to renegotiations only if

$$t_{\theta_D \theta_A} \in P_{\theta_D}, \quad (9)$$

for $\theta_D = l, h$. Therefore, the problem of the optimal renegotiation-proof mechanism may be written as $\max_{t_{\theta_D \theta_A}} EW(t_{\theta_D \theta_A})$, subject to (9), and incentive compatibility conditions. For the low-type D , the incentive compatibility condition may be written as

$$\begin{aligned} & \gamma V_D(t_{ll}; l) + (1 - \gamma) V_D(t_{lh}; l) \\ \geq & \gamma V_D(\Upsilon_\lambda^l(t_{hl}); l) + (1 - \gamma) V_D(\Upsilon_\lambda^l(t_{hh}); l). \end{aligned} \quad (10)$$

The left- (right-)hand side of 10 indicates the expected welfare of the low-type D if the state of the world is reported truthfully (untruthfully). To understand the expression on the right-hand side, note that if D_l misrepresents its type by announcing $\theta_D = h$, the outcome of the mechanism, $t_{h\theta_A}$, will not be Pareto efficient, in which case the outcome is renegotiated to $\Upsilon_\lambda^l(t_{h\theta_A})$ for $\theta_A = l, h$. The incentive compatibility condition for D_h may be obtained in a similar way, which yields

$$\begin{aligned} & \gamma V_D(\Upsilon_\lambda^h(t_{hh}); h) + (1 - \gamma) V_D(\Upsilon_\lambda^h(t_{hl}); h) \\ \geq & \gamma V_D(\Upsilon_\lambda^h(t_{lh}); h) + (1 - \gamma) V_D(\Upsilon_\lambda^h(t_{ll}); h). \end{aligned} \quad (11)$$

Letting DM denote the mechanism that maximizes (8) subject to the incentive compatibility and renegotiation proof constraints 9-11, we have

Lemma 2 *DM is robust to interim renegotiations.*

Figure 5: The Optimal *Direct Mechanism*, DM: t_l^{DM} , t_{hl}^{DM} , and t_{hh}^{DM} .

This Lemma states that regardless of the realized state of the world, at the interim stage the parties could not find an alternative mechanism that is jointly preferred to *DM*. A complete proof of this lemma is provided in Appendix B. To obtain an initial understanding of this result, note that the mechanism design problem faces essentially the same constraints in the interim and ex ante stages. The only difference is that *D* knows his type at the interim stage but not ex ante. But, *C*'s information about the state of the world at the interim stage will be identical to her prior at the ex ante stage. This makes it impossible to come up with an alternative mechanism that is preferred by both parties under some states of the world.

Letting $t_{\theta_D \theta_A}^{DM}$ denote the outcome of *DM* as a function of θ_D and θ_A , the following proposition establishes the properties of the optimal renegotiation-proof direct mechanism:

Proposition 2 *If $\gamma > \frac{1}{2}$, the outcome of the optimal renegotiation-proof direct mechanism, DM, satisfies the following conditions:*

- (i) $t_{ll}^{DM} = t_{lh}^{DM} \equiv t_l^{DM} \in P_l$.
- (ii) $t_{hh}^{DM}, t_{hl}^{DM} \in P_h$.
- (iii) *A low-type defending country is indifferent about truthfully revealing its type, i.e., condition (10) is satisfied with equality.*

To elaborate, the outcome of this direct mechanism is one of three tariff pairs, t_l^{DM} , t_{hh}^{DM} , or t_{hl}^{DM} , which is depicted in Figure (5). If the importing country announces $\theta_D = l$, the tariff pair $t_l^{DM} \in P_l$ will be implemented. If $\theta_D = h$ is announced, the outcome is either t_{hh}^{DM} or t_{hl}^{DM} depending on the signal received by the arbitrator. The importing country prefers t_{hh}^{DM} to t_{hl}^{DM} while the exporting country has the opposite preference.

Part (i) of this proposition implies that under the optimal mechanism, the DSB's signal will be redundant when the importing country reports a low political-economy pressure. To obtain an intuition for this result, note that the parties' joint welfare indicates risk aversion. Thus, other things equal, a certain outcome is preferred to a lottery.²⁴

Part (iii) of proposition 2 can be stated as

$$V_D(t_l^{DM}; l) = \gamma V_D(\Upsilon_\lambda^l(t_{hl}^{DM}); l) + (1 - \gamma) V_D(\Upsilon_\lambda^l(t_{hh}^{DM}); l). \quad (12)$$

This equality condition implies that the low-type defending country is indifferent between the tariff pair t_l^{DM} and a lottery between $\Upsilon_\lambda^l(t_{hl}^{DM})$ and $\Upsilon_\lambda^l(t_{hh}^{DM})$ with probabilities γ and $1 - \gamma$, respectively. Therefore, when renegotiations are possible, the optimal design of the agreement depends on the parameter of the bargaining process, λ . As will be discussed in more details in Subsections 7.3 and 7.5, this equation has important implications as it relates the optimal design of trade agreements to the existing informal norms or relative bargaining power of the parties in bilateral negotiations.

The benefit of incorporating arbitration in the agreement arises from the fact that a high-type defending country is more likely to receive a favorable policy recommendation than a low-type defending country. This increases the cost to a low-type D of mimicking a high type. To illustrate, consider an extreme case in which the arbitrator's signal is perfect (i.e., $\gamma = 1$). A perfectly informed arbitrator could sim-

²⁴Contrary to Part i of this Proposition, suppose that $t_{ll}^{DM} \neq t_{lh}^{DM}$ and consider a tariff pair, t' , that generates the same payoffs for the low-type D as does the lottery between t_{ll}^{DM} and t_{lh}^{DM} . Replacing t_{ll}^{DM} and t_{lh}^{DM} with t' in the mechanism, does not affect incentive compatibility condition, while generates a higher joint welfare.

ply assign the first-best outcome, namely $t^E(l)$ and $t^E(h)$. In the other extreme, an arbitrator with a completely uninformative signal, as in Beshkar (2010a), could be replaced with a purely randomizing device.

6 The Optimal *ALR*

The **ALR** provides an alternative to the direct mechanism (DM) that was introduced in Section 5. The two mechanisms are different only in the type of messages sent by the informed party and the set of outcomes that can be recommended and enforced by the arbitrator. These differences may be observed by comparing timelines of the DM and the **ALR** in Figures 1 and 4. First, under the **ALR**, the invocation of the escape clause, or lack thereof, is the message that is sent by the informed party at Date 2. That is different from a direct mechanism in which D 's message is chosen from its type space, $\theta \in \{l, h\}$.

The second difference between the **ALR** and a direct mechanism is the range of outcomes that can be recommended and enforced by the arbitrator. (Compare Date 3 in Figures 1 and 4.) As described in Definition 1, under the **ALR**, the arbitrator is bound to recommend the baseline tariff, τ_b , or the safeguard tariff, τ_s , and enforce any tariff pair that D chooses on the corresponding retaliation scheme.²⁵

My objective in this Section is to show that the outcome of the optimal direct mechanism (*DM*) may be implemented under the Arbitrated-Liability Regime. My approach, therefore, is to choose $t_b, t_s, r_l(\tau)$, and $r_h(\tau)$ under which the equilibrium of **ALR** replicates the outcome of the *DM*. The rest of this section provides the proof of the following proposition:

Proposition 3 *An **ALR** mechanism implements the outcome of DM if it satisfies*

²⁵In other words, the defending party has the right to choose any tariff, τ_D , above DSB's recommendation, while the DSB limits the magnitude of retaliation from the complaining party to $\tau_C < r_{\theta_A}(\tau_D)$.

the following conditions:

$$t_b = \Upsilon_\lambda^l(t_{hl}^{DM}), \quad (13)$$

$$\Upsilon_\lambda^l(t_s) = \Upsilon_\lambda(t_{hh}^{DM}), \quad (14)$$

$$\Upsilon_\lambda^l(\tau, r_{\theta_A}(\tau)) = \Upsilon_\lambda^l(t_{h\theta_A}^{DM}), \quad \forall \tau \leq \tau_{h\theta_A}^{DM}, \quad (15)$$

$$V_C(\tau, r_{\theta_A}(\tau)) = V_C(t_{h\theta_A}^{DM}), \quad \forall \tau \geq \tau_{h\theta_A}^{DM}, \quad (16)$$

for $\theta_A \in \{l, h\}$.

In the remainder of the paper, I use **ALR** to refer to the Arbitrated-Liability Regime that satisfies conditions 13-16. The mechanism characterized in this proposition is depicted in Figure (6). The baseline tariff in this figure satisfies condition 13, i.e., $t_b = \Upsilon_\lambda^l(t_{hl}^{DM})$.²⁶ That is, t_b is chosen such that when $\theta = l$, ex post renegotiations from t_{hl}^{DM} lead to t_b .

Similarly, condition 14 implies that the safeguard tariff pair, t_s , is chosen such that when $\theta = l$, renegotiations from t_s and t_{hh}^{DM} lead to the same outcome. Moreover, as required by the definition of the **ALR**, t_b and t_s in Figure (6) specify the same tariff for C , i.e., $r_s = r_b$.

Condition 15 requires the retaliation schemes, $r_l(\tau)$ and $r_h(\tau)$, to be the ‘bargaining paths’ that correspond to t_{hl}^{DM} and t_{hh}^{DM} , respectively, when $\theta = l$. These retaliation schemes ensure that a low-type D will be in the same bargaining position regardless of the tariff that it chooses above the arbitrator’s recommended level.

To clarify further, I discuss each of these retaliation schemes separately. For any τ in the $[\tau_b, \tau_{hl}^{DM}]$ interval, $r_l(\tau)$ characterizes the set of tariff pairs that would be renegotiated to $t_b = \Upsilon_\lambda^l(t_{hl}^{DM})$ in the ex post stage if $\theta = l$. For $\tau > \tau_{hl}^{DM}$, $r_l(\tau)$ authorizes a sufficiently large retaliation that preserves C ’s payoffs at t_{hl}^{DM} , namely:

$$V_C(\tau, r_l(\tau)) = V_C(t_{hl}^{DM}), \quad \forall \tau \geq \tau_{hl}^{DM}. \quad (17)$$

This latter part of the retaliation scheme, which coincides with C ’s indifference curve

²⁶Recall that $\Upsilon_\lambda^\theta(t)$ was defined by equations 4 and 5.

The retaliation scheme $r_h(\tau)$ has a similar interpretation. In particular, for any $\tau_s \leq \tau \leq \tau_{hh}^{DM}$, $(\tau, r_h(\tau))$ is a tariff pair that will be renegotiated to $\Upsilon_\lambda^l(t_{hh}^{DM})$ if $\theta = l$. For $\tau > \tau_{hh}^{DM}$, $(\tau, r_h(\tau))$ is a tariff pair that preserves C 's payoffs at t_{hh}^{DM} . Finally, note that the safeguard tariff pair, t_s , defined by (14), is located on the $r_h(\tau)$ schedule.

Lemma 3 *Given arbitrator's judgment, θ_A , the outcome of ex post renegotiations under the **ALR** is $t_{h\theta_A}^{DM}$ iff $\theta = h$ and $\Upsilon_\lambda^l(t_{h\theta_A})$ iff $\theta = l$.*

Proposition 4 *If the ALR mechanism is in place, C proposes $t^S = t_l^{DM}$ as an alternative agreement in the interim stage, and D will accept (reject) this proposal if $\theta = l$ ($\theta = h$).*

Proof. In the interim renegotiations, a low-type D will accept a proposal, t^S , if and only if

$$V_D(t^S; l) \geq \gamma V_D(t_b; l) + (1 - \gamma) V_D(\Upsilon_\lambda^l(t_s); l). \quad (18)$$

On the other hand, we know from 12 that

$$V_D(t_l^{DM}; l) = \gamma V_D(\Upsilon_\lambda^l(t_{hl}^{DM}); l) + (1 - \gamma) V_D(\Upsilon_\lambda^l(t_{hh}^{DM}); l). \quad (19)$$

Therefore, given that $t_b = \Upsilon_\lambda^l(t_{hl}^{DM})$ and $\Upsilon_\lambda^l(t_s) = \Upsilon_\lambda^l(t_{hh}^{DM})$ due to 13 and 14, $t^S = t_l^{DM}$ will satisfy 18 with equality.

Now consider the incentives of a high-type D regarding a proposed tariff pair t^S . A high-type D will accept t^S if and only if

$$V_D(t^S; h) \geq \gamma V_D(t_{hh}; h) + (1 - \gamma) V_D(t_{hl}; h). \quad (20)$$

However, the incentive compatibility of the DM for the high-type D implies that

$$V_D(t_l^{DM}; h) < \gamma V_D(t_{hh}^{DM}; h) + (1 - \gamma) V_D(t_{hl}^{DM}; h). \quad (21)$$

Therefore, the high-type D will reject the settlement proposal $t^S = t_l^{DM}$.

Finally consider C 's incentive to propose $t^S = t_l^{DM}$. Given that this proposal is only accepted by a low-type D , C will be willing to propose such a tariff pair if and only if

$$V_C(t_l^{DM}) > \gamma V_C(\Upsilon_\lambda^l(t_{hl}^{DM})) + (1 - \gamma) V_C(\Upsilon_\lambda^l(t_{hh}^{DM})). \quad (22)$$

If this condition is satisfied, then together with condition 19, it implies that

$$W(t_l^{DM}; l) > \gamma W(\Upsilon_\lambda^l(t_{hl}^{DM}); l) + (1 - \gamma) W(\Upsilon_\lambda^l(t_{hh}^{DM}), l).$$

In other words, when $\theta = l$, condition 22 would require that the parties' expected joint welfare be higher under the lottery between $\Upsilon_\lambda^l(t_{hl}^{DM})$ and $\Upsilon_\lambda^l(t_{hh}^{DM})$ with probabilities γ and $1 - \gamma$, respectively, than under t_l^{DM} .

Therefore, if $t^S = t_l^{DM}$ in the interim stage of the **ALR** game, the direct mecha-

nism DM is implemented.

Now suppose that C could propose a mechanism, DM' , to D that results in a higher expected payoffs for C than offering $t^S = t_l^{DM}$. Since D has the option to go with DM , D 's payoff under DM' should be at least as high as his expected payoffs under DM . Therefore, DM' would be acceptable only if it leads to higher expected joint-welfare than the DM . But this is contrary to the assumption that DM maximizes the expected joint welfare of the parties subject to incentive compatibility and renegotiation proof conditions. ■

Lemma 3 and Proposition 4 together imply that, as depicted in Figure (6), the outcome of the **ALR** is one of three tariff pairs, t_l^{DM} , t_{hl}^{DM} , and t_{hh}^{DM} , which replicates the optimal direct mechanism, DM . This completes the proof of Proposition 3.

We, therefore, have shown that the optimal direct mechanism, DM , which does not look ‘practical’, has a simple representation, i.e., the **ALR**, which resembles the customary institutional structure of international organizations.

7 Implications of the Arbitrated-Liability Regime

I will now explore different implications of the optimal trade/arbitration agreement that was characterized in the previous sections. I state the results of this section informally as they are intuitive given the formal results that we have obtained so far.

7.1 Depth of liberalization, Magnitude of Escape, and DSB's Monitoring Quality

Transition from GATT to the WTO included a notable reform in the escape clause. Article XIX of GATT indicates that an exporting country that is affected by a safeguard measure (or escape) could suspend substantially equivalent concessions as a means of receiving compensation. Under the WTO's Agreement on Safeguards, however, the affected exporting countries' right to receive compensation is subject to DSB's authorization. As a result of reforms introduced by the Agreement on Safeguards, an importing country that adopts a safeguard measure is not required

to compensate affected countries for a period of four years, unless the DSB rules the adopted safeguards illegal. As described by Pelc (2009), over time “compensation after escape” has been largely replaced with “appeal to exception” in the GATT/WTO.²⁷

My model suggests that these reforms in the safeguard rules might have been prompted by an increase in the accuracy of the DSB in monitoring and verifying different trade-related contingencies. As discussed at the end of Section 5, with a fully-informed arbitrator, the optimal agreement achieves the first best outcome. The first-best outcome includes a safeguard clause that allows a relatively large escape from the baseline tariff concession with no compensation requirement. Moreover, under the first-best outcome the parties fully comply with the DSB’s ruling and no retaliation will take place. Finally, as the DSB’s signal improves, the optimal baseline tariff decreases. In summary,

Remark 1 *As DSB’s signal improves, the optimal trade agreement features more aggressive trade liberalization (i.e., a lower tariff binding τ_b), and a greater magnitude of escape (i.e., a greater $\tau_s - \tau_b$).*²⁸

7.2 Early Settlement, Non-Compliance, and DSB’s Biased Ruling Pattern

Under the **ALR**, a dispute will arise if an importing country wants to apply a tariff above the committed binding. The model predicts that when the political-economy parameter has a low realization, the two parties (i.e., the importer and the exporter) will find a mutually accepted solution (namely, an early settlement) and arbitration will not be invoked. The early settlement agreement involves a tariff pair, t_t^{DM} , that is more favorable to the importing country than the agreement’s baseline tariff, t_b . This implies that under this regime, the exporting country will tolerate small deviations

²⁷Maggi and Staiger (2012) interpret this change in the escape rules as an evolution from liability rule (which requires compensation for breach) to property rule (which requires a consensus for modification of concessions).

²⁸In Maggi and Staiger (2012), an increase in DSB’s signal quality increases the efficiency of a property-rule system over a liability-rule system.

from the agreement by the importing country. The source of this *forbearance* is the fact that the arbitration system is imperfect and adjudication may result in a worse situation for the exporting country.^{29,30}

Non-compliance with the DSB’s ruling and retaliatory actions also occur in equilibrium. The model predicts that an importing country that is under high political-economy pressures will *always* decline to limit its tariff to the level that is determined by the DSB on the equilibrium path. However, the level of retaliations that such an importing country will face, depends on the DSB’s findings regarding the legitimacy of a contingent protection measure.

This sharp prediction about non-compliance with the DSB’s ruling could inform the current debate on the rulings of the WTO’s dispute settlement body since its inception in 1995. The data on the official rulings of the DSB reveals a high disparity between the success rates of the complaining and defending parties. As reported by Colares (2009), the DSB rules against the defending party in more than 88 percent of cases where the subject of dispute is related to trade remedies.³¹ In some categories of disputes this disparity is even more dramatic. For example, in litigations regarding the safeguard measures adopted to protect domestic industries against potentially harmful surge in imports, the DSB has always ruled against the defending party (Sykes 2003).

Some observers have assessed this pro-complainant ruling pattern as unsatisfactory. For example, Sykes (2003) and Grossman and Sykes (2007) argue that the DSB’s interpretation of the WTO Agreement has made it increasingly difficult for the governments to resort to the escape clause, which frustrates the purpose of the WTO Agreement on Safeguards. Colares (2009) attributes the DSB’s bias to the normative views of the individuals who are involved in the DSB and argues that the asymmetrical pattern of the DSB’s ruling is “the result of a process of authorita-

²⁹In practice, many safeguard measures are not formally challenged in the WTO. Such cases may reflect the forbearance predicted by this paper.

³⁰Within a repeated-game framework, Bowen (2011) provides a model in which signatories of a trade agreement show *forbearance*, where one country withholds retaliation when its trading partner receives a shock. The forbearance under the **ALR** is different in that the exporting country forgoes *its right to challenge* the importing country’s illegal measure in the dispute settlement process.

³¹For non-trade remedy cases this rate is 83.33%.

tive normative evolution (i.e., rule development) that has expressed itself with a tilt favoring complainants.”

The results of this paper suggest that an optimally-designed WTO would show some bias towards the complainants in the trade disputes.³² It can be observed from Figure (6) that even if the DSB finds evidence in favor of the defending country, i.e., $\theta_A = h$, the defending country will not have a full victory. In such a case, the DSB’s optimal ruling is to allow the defending country to increase its tariff from τ_b to τ_s with impunity. But when $\theta = h$, the defending country will violate the ruling of the DSB by setting $\tau = \tau_{hh} > \tau_s$. In other words, even if the DSB finds evidence in favor of the defendant, the level of protection adopted by the defending party will exceed the level that is authorized by the DSB.

7.3 The Role of Informal Norms or Bargaining Power

The introduction of the dispute settlement process to the WTO has been interpreted as moving away from a power-based renegotiation system towards a rules-based system. In a power-based relationship, parties engage in bargaining without being constrained by any principle. In contrast, a rules-based institution has pre-specified rules for negotiating or renegotiating the terms of the agreement. A well-known limitation of the rules-based approach is that parties may mutually agree to ignore the rules.

In this subsection I demonstrate how the optimal choice of a rules-based system is constrained by the existing informal norms of renegotiation among parties. In particular, I show that the optimal *formal* rule of renegotiation is a function of the relative bargaining power between the parties. This point can be understood from the graphical illustration of the optimal **ALR** in Figure (6). Recall that the optimal retaliation scheme, given by menus $r_l(\tau)$ and $r_h(\tau)$ in Figure (6), are the *bargaining paths* as defined by equation 15. These paths reflect the bargaining process between the parties when the true state of the world is $\theta = l$, such that any tariff combinations

³²This result, however, does not rule out the possibility that the DSB may be biased too much in favor of the complaining parties, as suggested by Sykes (2003) and Grossman and Sykes (2007).

on each of these paths would be renegotiated to the same tariff pair on the set of Pareto efficient tariffs (P_l) , respectively.

It is instructional to examine the optimal retaliation rule for two extreme cases where the defending country has no bargaining power ($\lambda = 0$) and when the defending country has all the bargaining power ($\lambda = 1$) in bilateral settlement negotiations with the exporting country. If the defending country has no bargaining power ($\lambda = 0$), the lower part of menus $r_l(\tau)$ and $r_h(\tau)$ in Figure (6) will coincide with the defending country's indifference curves. In particular, $r_l(\tau)$ ($r_h(\tau)$) will be the upper envelope of the indifference curves of D_l and C that go through t_{hl}^{DM} (t_{hh}^{DM}).

Therefore, when $\lambda = 0$, the punishment for non-compliance is just enough to make the low-type defending country indifferent between compliance and non-compliance with the DSB ruling.³³ This result follows due to the fact that retaliations are efficiency-reducing and, thus, it is optimal to choose the lowest amount of punishment that constitutes an incentive-compatible mechanism.

In the other extreme of relative bargaining power, $\lambda = 1$, the retaliation menus $r_l(\tau)$ and $r_h(\tau)$ coincide with the indifference curves of the complaining country. That is because when the defending country has all the bargaining power, the bargaining path consists of the exporting country's indifference curve. In general,

Remark 2 *Under an optimal agreement, the higher is the bargaining power of the defending country in renegotiations, the more severe retaliations must be authorized against noncompliance. If $\lambda = 1$, the retaliations fully compensate the complaining country under an optimal agreement. If $\lambda < 1$, the magnitude of retaliations are not sufficient to fully compensate the complaining country. If $\lambda = 0$, retaliations are just enough to make a low-type defending country indifferent about truth-telling.*

Intuitively, a more powerful defector must be subject to harsher punishment so as to preclude any benefit to the defector from deviating to an off-equilibrium point along the punishment scheme and then offering a Pareto-improving proposal. In other words, any point on the punishment scheme should put the defector in a

³³It is also worth noting that this punishment scheme compensates the exporting country only for a fraction of its loss due to importing country's increased protection.

sufficiently bad bargaining position such that renegotiation from that point would lead to an outcome that is (weakly) dominated by the on-equilibrium outcome of the agreement.

7.4 The Property Rule

The Property Rule is another important legal principle that is often used to regulate the exchange of entitlements. Under this rule, both parties' consent is required to change the default entitlements. In an international trade cooperation setting, the property rule may be interpreted as allocating the right of market access to each country and letting governments to renegotiate those entitlements based on mutual consent.

An important difference between the liability and the property rules, therefore, is that the latter leaves renegotiations unrestricted. In other words, the property rule system is akin to a power-based dispute settlement procedure in which the outcome of negotiations is determined directly by the relative bargaining power of the parties. The analysis of this paper, however, shows that when efficient side payments are unavailable, a power-based relationship does not necessarily lead to an efficient outcome. In particular, as was elaborated in Remark 2, in general under an optimal escape rule, the exporting country is not compensated fully for its loss. In contrast, under the property rule the affected exporting country would be more-than-compensated for its loss by sharing the rent from increased protection in the importing country. We can, thus, conclude that:

Remark 3 *The Property Rule is suboptimal as an escape provision.*

Although the property rule is irreconcilable with efficient breach (i.e., escape), it can be part of an optimal mechanism if instead of giving the *right of market access* to the exporting countries, we give the *right of import protection* to the importing countries. In fact, the *DM* mechanism as depicted in Figure (6), can be immediately interpreted as a Property Rule system in which the DSB's recommended outcome is either t_{hl}^{DM} (when $\theta_A = l$) or t_{hh}^{DM} (when $\theta_A = h$). Just as in *DM*, any deviation

from these recommendations will require both parties' consent. In that case, if $\theta = h$, then the parties will implement the DSB's recommendation (since it is Pareto efficient given $\theta = h$), and if $\theta = l$, then ex post renegotiations lead parties to the tariff pair $\Upsilon_{\lambda}^l(t_{h\theta_A}^{DM})$.³⁴ Therefore,

Remark 4 *The optimal agreement under the property-rule regime characterizes higher default tariff commitments than under the liability-rule regime. Moreover, under the property-rule regime, the agreement tariffs are renegotiated down whenever the political-economy parameter in the importing country is low.*

In other words, under the property-rule regime, the parties do not commit to substantial tariff cuts ex ante, while tariff cuts are negotiated in each period. The actual trade agreements, however, are structured in a starkly different way, such that the default tariffs are set low and parties are given flexibility via an escape clause. This observation suggests that within the common structure of tariff agreements, the property rule is not the optimal form of remedy.

The above results about liability and property rules are in contrast to Maggi and Staiger's (2012), in which property rule may dominate liability rule as an escape rule. In Maggi and Staiger (2012), a property rule becomes relatively more efficient as the quality of the DSB's signal improves. My model, on the other hand, suggests that these rules could achieve the same outcome in the equilibrium. Nevertheless, the structure of an optimal liability-rule system bears a closer resemblance to the structure of the actual international trade institutions.

7.5 The Reciprocity Principle

The principle of reciprocity has been viewed as one of the pillars of the GATT/WTO. Therefore, it is of interest to examine the optimality of the reciprocity principle as a

³⁴To be concrete, consider the game that is induced by the property-rule regime that I just described. In the interim stage, i.e., after the realization of the state of the world but before playing the mechanism, the parties can negotiate a settlement to avoid arbitration. This game is identical to the game induced by the arbitrated-liability regime (*ALR*). Therefore, as was shown in Proposition 3, *DM* is implementable as a Perfect Bayesian Equilibrium of this game.

remedy rule. To this end, consider a definition of the reciprocity principle introduced by Bagwell and Staiger (1999): “A set of tariff changes conform to the principle of reciprocity if it brings about equal changes in the import volume of the two countries at the initial prices.” With this definition, the reciprocity principle cannot be an optimal retaliation rule since it is not a function of the relative bargaining power of the parties.

Another interesting question is whether the reciprocity principle prescribes too much or too little retaliation. As suggested by Remark 2, the answer to this question depends critically on the relative bargaining power of the governments in renegotiations. Recall that when $\lambda = 0$, the optimal retaliation is just enough to make the low-type defending country indifferent about compliance with the DSB’s ruling. Under the reciprocity principle, however, the low type defending country is strictly worse off by non-compliance. Therefore, the punishment associated with reciprocal retaliations is greater than the optimal punishment when $\lambda = 0$.³⁵

On the other hand, when $\lambda = 1$, the magnitude of optimal retaliation is sufficiently large that fully compensates the exporting country for its loss due to deviation of the defending country. Reciprocal retaliations, however, compensate the exporting country only for a fraction of its loss.³⁶ Therefore,

Remark 5 *If a defending country has a sufficiently low (high) bargaining power in the dispute settlement process, the magnitude of optimal retaliation is less (more) than reciprocal.*

8 Conclusion

This paper shows that the optimal trade agreement takes a simple form that resembles a safeguard arrangement with a compensation mechanism: a baseline tariff, a

³⁵This is similar to the finding of Beshkar (2010a) that punishment under an optimal agreement is barely enough to prevent inefficient breach.

³⁶As explained by Bagwell (2008), this result is consistent with the traditional interpretation of the liability rule, which prescribes a compensation level that leaves the breached-upon party unaffected.

safeguard tariff that must be approved by the DSB, and a retaliation scheme that determines the size of retaliations in case of noncompliance. It was shown that the optimal ruling by the DSB demonstrates a bias in favor of the complaining (i.e., exporting) country, such that in any dispute between the parties, the DSB should rule at least partly against the proposed increase in trade protection by the defending country.

This study identifies a novel limitation of international trade law and provides guidelines for the design of an optimal remedy mechanism for an international agreement. It is demonstrated that the design of a rules-based system of dispute settlement is constrained by informal norms and power relations between the disputing parties. In particular, if it was possible to completely isolate the dispute settlement procedure from informal negotiations, the optimal level of retaliations would be smaller. On the other hand, if the dispute settlement negotiations are unrestricted by any rules, the size of equilibrium retaliations will be higher. This reduction in the equilibrium level of punishment is what a rules-based dispute settlement procedure can achieve.

The model of this paper provides a sharp prediction regarding the pattern of early settlement, litigation, and non-compliance with DSB rulings, such that on the equilibrium path we will always observe non-compliance with the DSB ruling if the early settlement negotiations fail.

Richer results may be obtained by extending this model to capture a more complicated informational structure. For example, the analysis in this paper was simplified by ruling out the possibility that the uninformed disputing party may receive a noisy signal of the true state of the world. In practice, however, the disputing parties may have better knowledge and cheaper ways to acquire information about the prevailing state of the world than the arbitrators. The proposed model can be extended by allowing for private signaling, in which the uninformed parties could privately observe a noisy signal about the prevailing state of the world during the pre-arbitration negotiations.

I modeled tariffs as ‘public actions’ that are externally enforced. Although modeling verifiable actions as public actions has been a common practice in the mechanism design literature, Watson (2007) and Buzard and Watson (2012) show that there are

important reasons to be cautious about this *simplifying* assumption in some economic applications. In particular, if the opportunity to take an action is nondurable in a hold-up problem, assuming that a central planner could choose these actions is not a reasonable characterization of the real world. Moreover, such an assumption is not innocuous as it changes the set of implementable value functions. But since tariffs are essentially reversible, it may be justifiable to treat them as public actions. For example, according to Article 6 of the Safeguard Agreement, in case of provisional safeguard measures, tariff increases must be “*promptly refunded if the subsequent investigation ... does not determine that increased imports have caused or threatened to cause serious injury to a domestic industry.*” Nevertheless, modeling tariffs as ‘individual actions’ could be illuminating since the negative economic impact of tariff increases may not be completely eliminated by the promise of future refunds.³⁷

Another important area for future research is to consider countries that are asymmetric in technological or political-economy parameters. Such an extension of the model would be particularly useful in understanding whether all types of countries should be equal before the WTO law or whether it is optimal to apply different standards for various types of countries, such as smaller developing countries.

A limitation of the current study is the assumption that the mechanism designer is fully aware of the informal norms and the power relation between the disputing parties. Without the knowledge about how parties will split the gains from renegotiations, we can only determine a sufficient condition for renegotiation-proofness of the mechanism as in Neeman and Pavlov (2013). Nevertheless, the proposed framework reveals the limit of a rules-based system even when the arbitrator is fully aware of the details of informal relations between the parties. One potentially interesting extension of the current set up is to introduce imperfection in the mechanism designer’s knowledge about the informal renegotiation procedure.

The application of this model is not limited to the DSB. This paper may be

³⁷In the case of Generalized System of Preferences in the US, Hakobyan (2013) provides evidence that a tariff increase has a negative impact on trade volume even when there is an expectation that such tariffs will be refunded in the future.

applied to the analysis of any third-party arbitrator with informative insights about the dispute. For example experts from World Health Organization and International Monetary Fund can play a useful role in the arbitration process in cases related to health and exchange rate policy, respectively.³⁸

Appendix A. Construction of the Welfare Functions

Markets

Consider a pair of distinct goods m and x with demand functions in the home country (no $*$) and the foreign country ($*$) given by:

$$\begin{aligned} D_m(p_m) &= 1 - p_m, & D_x(p_x) &= 1 - p_x, \\ D_m^*(p_m^*) &= 1 - p_m^*, & D_x^*(p_x^*) &= 1 - p_x^*, \end{aligned} \quad (23)$$

where p (with the appropriate index) represents the price of a good in a certain country. Specific import tariffs, τ and r , that are chosen by countries as the only trade policy instrument, create a gap between domestic and foreign prices. In particular, $p_m = p_m^* + \tau$ and $p_x = p_x^* - r$.

Both countries produce both goods using the following supply functions:

$$\begin{aligned} Q_m(p_m) &= p_m, & Q_x(p_x) &= bp_x, \\ Q_m^*(p_m^*) &= bp_m^*, & Q_x^*(p_x^*) &= p_x^*. \end{aligned} \quad (24)$$

Assuming $b > 1$, the home country will be a natural importer of m and a natural exporter of x .

World market clearing condition for good m is $D_m(p_m) + D_m^*(p_m - \tau) = Q_m(p_m) + Q_m^*(p_m - \tau)$. Substituting for the supply and demand functions from (23) and (24), the market clearing condition can be rewritten as $2 - 2p_m + \tau = p_m + b(p_m - \tau)$. Solving for p_m yields $p_m = \frac{2 + (1+b)\tau}{3+b}$. Similarly, using the world market clearing condition for good x , the home market price for good x can be calculated; $p_x = \frac{2(1-r)}{3+b}$.

Components of Welfare

Under this model, the market-clearing price of m (x) depends only on the home (foreign) tariff. Let $p_m(\tau)$ and $p_x(r)$ respectively denote the equilibrium prices of m and x in the home country. If import tariffs are non-prohibitive (i.e., if they are sufficiently small) trade occurs between the countries and the home consumers'

³⁸I am grateful to Helen Milner for pointing this out to me.

surplus from the consumption of m and x will be given, respectively, by

$$\begin{aligned}\psi_m(\tau) &\equiv \int_{p_m(\tau)}^1 D_m(u) du = \frac{1}{2} - p_m + \frac{1}{2}p_m^2 = \frac{1}{2} \left(\frac{(1+b)(1-\tau)}{3+b} \right)^2, \\ \psi_x(r) &\equiv \int_{p_x(r)}^1 D_x(u) du = \frac{1}{2} \left(\frac{1+b+2r}{3+b} \right)^2.\end{aligned}$$

Moreover, the home producers' surplus from the sale of m and x will be given by

$$\begin{aligned}\pi_m(\tau) &\equiv \int_0^{p_m(\tau)} Q_m(u) du = \frac{1}{2}p_m^2 = \frac{1}{2} \left(\frac{2+(1+b)\tau}{3+b} \right)^2, \\ \pi_x(r) &\equiv \int_0^{p_x(r)} Q_x(u) du = \frac{1}{2}bp_x^2 = 2b \left(\frac{1-r}{3+b} \right)^2.\end{aligned}$$

The government's tariff revenue is given by

$$R(\tau) \equiv \tau M_m(p_m(\tau)) = \frac{(b-1)\tau - 2(1+b)\tau^2}{3+b}.$$

where $D_m(p_m)$, the import demand for good m in the home country, is given by

$$D_m(p_m) \equiv D_m(p_m) - Q_m(p_m) = 1 - 2p_m = \frac{b-1-2(1+b)\tau}{3+b}.$$

Politically weighted welfare from the importing sector in home country is given by

$$\begin{aligned}u(\tau; \theta) &\equiv \psi_m(\tau) + \theta\pi_m(\tau) + R(\tau) \\ &= \frac{1}{(3+b)^2} \left\{ \frac{1}{2}(1+b)^2 + 2\theta + [2\theta(1+b) - 4]\tau + \left[\frac{1+\theta}{2}(1+b)^2 - 2(3+b)(1+b) \right] \tau^2 \right\}.\end{aligned}\tag{25}$$

The home government's welfare from the exporting sector is a function of foreign tariff, τ^* :

$$v(\tau^*) \equiv \psi_x(\tau^*) + \pi_x(\tau^*) = \frac{1}{(3+b)^2} \left\{ \frac{(1+b)^2}{2} + 2b + 2(1-b)\tau^* + 2(1+b)(\tau^*)^2 \right\}.\tag{26}$$

Now using the functions u and v constructed above, we can define welfare of the defending (D) and complaining (C) countries. Letting τ and r denote the tariffs of

D and C respectively, the welfare functions of the governments are

$$\begin{aligned} V_D(t; \theta) &\equiv u_D(\tau; \theta) + v_D(r), \\ V_C(t) &\equiv u_C(r; l) + v_C(\tau). \end{aligned}$$

Note that given the symmetry of payoff functions in the two countries, we can drop the country subscripts from u and v functions.

Appendix B. Proofs

Lemma 4 *For any $\alpha \in (0, 1)$ and $\tau_1, \tau_2, \tau_3 < \tau^N(\theta)$, if $u(\tau_3; \theta) = \alpha u(\tau_1; \theta) + (1 - \alpha) u(\tau_2; \theta)$, then $v(\tau_3) > \alpha v(\tau_1) + (1 - \alpha) v(\tau_2)$.*

Proof. As shown in (25) and (26), u and v are quadratic functions that may be written as

$$\begin{aligned} u(\tau; \theta) &= -A\tau^2 + B\tau + F, \\ v(\tau) &= C\tau^2 - D\tau + G, \end{aligned}$$

where

$$A > C > 0, \quad B > D > 0, \quad AD - BC > 0. \quad (27)$$

The first set of inequalities follows since the importer's and the joint welfare functions are concave and the exporter's welfare is convex. The second set of inequalities are satisfied because the importer's and the joint welfare functions are increasing in τ for $\tau \rightarrow 0$ and the exporter's welfare is decreasing in τ . The last inequality is satisfied since $\tau^N = \frac{B}{2A}$, $\tau^E = \frac{B-D}{2(A-C)}$, and $t^N(\theta) > t^E(\theta)$.

Given this formulation, the Arrow-Pratt measure of risk aversion for u and $-v$ are respectively given by

$$-\frac{u''(\tau; \theta)}{u'(\tau; \theta)} = -\frac{-2A}{-2A\tau + B},$$

and

$$-\frac{v''(\tau)}{v'(\tau)} = -\frac{-2C}{-2C\tau + D},$$

According to the Arrow-Pratt theorem, since u and $-v$ are concave, if $-\frac{u''(\tau;\theta)}{u'(\tau;\theta)} > -\frac{v''(\tau)}{v'(\tau)}$ then the certainty equivalent of u is always smaller than the certainty equivalent of $-v$ for any probability distribution of τ . Therefore, to prove the lemma, it is sufficient to show that $-\frac{u''(\tau;\theta)}{u'(\tau;\theta)} > -\frac{v''(\tau)}{v'(\tau)}$, or

$$\begin{aligned} -\frac{-2A}{-2A\tau + B} &> -\frac{-2C}{-2C\tau + D} \\ \iff \frac{A}{-2A\tau + B} &> \frac{C}{-2C\tau + D}. \end{aligned}$$

For $\tau < \tau^N$ we have $-2A\tau + B > 0$ and $-2C\tau + D > 0$, thus $-\frac{u''(\tau;\theta)}{u'(\tau;\theta)} < -\frac{v''(\tau)}{v'(\tau)}$ iff

$$A(-2C\tau + D) > C(-2A\tau + B),$$

or iff

$$AD - BC > 0.$$

The last inequality is satisfied according to (27). QED ■

Lemma 5 *For any $t_1, t_2 \in P_\theta$ and $\alpha \in (0, 1)$, there exists $t_3 \in P_\theta$ such that*

$$V_D(t_3; \theta) = \alpha V_D(t_1; \theta) + (1 - \alpha) V_D(t_2; \theta), \quad (28)$$

and

$$V_C(t_3) > \alpha V_C(t_1) + (1 - \alpha) V_C(t_2). \quad (29)$$

Proof. Since $t_1, t_2 \in P_\theta$ and $t_1 \neq t_2$, we must have $\tau_1 \neq \tau_2$ and $r_1 \neq r_2$. Then, according to Lemma 4, there must exist τ' and r' such that

$$u(\tau', \theta) = \alpha u(\tau_1; \theta) + (1 - \alpha) u(\tau_2; \theta), \quad (30)$$

$$v(\tau') > \alpha v(\tau_1) + (1 - \alpha) v(\tau_2). \quad (31)$$

and

$$u(r', l) = \alpha u(r_1; l) + (1 - \alpha) u(r_2; l), \quad (32)$$

$$v(r') > \alpha v(r_1) + (1 - \alpha) v(r_2). \quad (33)$$

Relationships (30) and (33) imply that

$$V_D(t', \theta) > \alpha V_D(t_1; \theta) + (1 - \alpha) V_D(t_2; \theta), \quad (34)$$

Similarly, relationships (31) and (32) imply that

$$V_C(t') > \alpha V_C(t_1) + (1 - \alpha) V_C(t_2). \quad (35)$$

If $t' \notin P_\theta$, then there must exist $t'' \in P_\theta$ such that $V_D(t'', \theta) > V_D(t', \theta)$, and $V_C(t'') > V_C(t')$. Therefore, conditions (34) and (35) imply

$$V_D(t'', \theta) > \alpha V_D(t_1; \theta) + (1 - \alpha) V_D(t_2; \theta), \quad (36)$$

and

$$V_C(t'') > \alpha V_C(t_1) + (1 - \alpha) V_C(t_2). \quad (37)$$

Now defining t_3 such that $t_3 \in P_\theta$ and

$$V(t_3, \theta) = \alpha V_D(t_1; \theta) + (1 - \alpha) V_D(t_2; \theta), \quad (38)$$

condition (36) implies that $V_D(t_3, \theta) < V_D(t'', \theta)$. Then, since $t'', t_3 \in P_\theta$, we must have

$$\tau' < \tau_3 \text{ and } r' > r_3,$$

which in turn implies that

$$V_C(t_3) > V_C(t''). \quad (39)$$

Therefore, $t_3 \in P_\theta$ satisfies conditions (28) and (29). QED. ■

Proof of Proposition 2. Parts (i) and (ii) To show that $t_u^{DM} = t_{lh}^{DM}$, suppose, on the contrary, that $t_u^{DM} \neq t_{lh}^{DM}$. Then according to Lemma 5, since $t_{lh}^{DM}, t_u^{DM} \in P_l$, there exists $t' \in P_l$ such that

$$V_D(t', l) = \gamma V_D(t_u^{DM}; l) + (1 - \gamma) V_D(t_{lh}^{DM}; l), \quad (40)$$

and

$$V_C(t') > \gamma V_C(t_u^{DM}) + (1 - \gamma) V_C(t_{lh}^{DM}). \quad (41)$$

To prove $t_u^{DM} = t_{lh}^{DM} \equiv t_l^{DM}$, it is sufficient to show that the mechanism that is obtained by replacing t_{lh}^{DM} and t_u^{DM} with t' satisfies the incentive compatibility conditions and generates a higher expected joint payoff than that of the DM (hence, the contradiction).

First note that condition (41) implies that C prefers t' to playing DM . Moreover, given (40), replacing t_{lh}^{DM} and t_u^{DM} with t' will not impact the welfare of the low-type defending country regardless of its announcement. Therefore, the incentive compatibility condition (10) will be still satisfied if we replace t_{lh}^{DM} and t_u^{DM} with t' . Finally, the incentive compatibility condition of the high-type D (i.e., condition

11) may or may not be satisfied. In either case the expected joint welfare under this alternative mechanism exceeds that of DM . This is a contradiction as DM is optimal. Thus, $t_{ll}^{DM} = t_{lh}^{DM}$.

Part (iii) By way of contradiction, suppose that the incentive compatibility condition for a low-type D (i.e., condition 10) is not binding. This means that t_{hh}^{DM} and t_{hl}^{DM} could be shifted further towards $t^E(h)$ without upsetting this condition (see figure 5). Such an adjustment in t_{hh}^{DM} and t_{hl}^{DM} will increase the joint welfare of the parties when $\theta = h$, while it has no impact on the welfare when $\theta = l$. Thus, the expected joint welfare could be improved if condition (10) is not binding. ■

Proof of Lemma 2. Consider C 's problem of offering an alternative mechanism at Date 1-1 when the current (status quo) mechanism is DM . Note that although D has realized its type at this point, C 's information about the state of the world is identical to his information at the ex ante stage (i.e., before Date 0).

The DM may be renegotiated successfully only if there is an alternative mechanism that is incentive compatible and renegotiation proof and weakly preferred by both parties. To prove that there is no such alternative to DM at the interim stage, I follow Holmström and Myerson's (1983) approach by considering the possibility that both parties would prefer an alternative mechanism under each (private) realization of the state of the world.

Consider an alternative agreement denoted by $DM' \equiv (t'_{ll}, t'_{lh}, t'_{hl}, t'_{hh})$. There is a positive chance that this alternative agreement is accepted unanimously if and only if one of the following three cases is possible:

Case 1: The alternative agreement increases the expected payoffs of every type of every player. This is impossible because it simply means that the status quo mechanism (i.e., the DM) is not optimal.

Case 2: Only the high-type D and C vote for mechanism DM' . This case prevails if the following four conditions are satisfied:

1) The high-type D prefers DM' to DM :

$$\gamma V_D(t'_{hh}; h) + (1 - \gamma) V_D(t'_{hh}; h) \geq \gamma V_D(t_{hh}^{DM}; h) + (1 - \gamma) V_D(t_{hh}^{DM}; h). \quad (42)$$

2) The low-type D prefers DM to DM' if it reveals its type truthfully under DM' :

$$\gamma V_D(t'_{ll}; l) + (1 - \gamma) V_D(t'_{lh}; l) \leq V_D(t_l^{DM}; l), \quad (43)$$

3) The low-type D prefers DM to DM' if it reveals its type untruthfully under DM' :

$$\gamma V_D(t'_{hl}; l) + (1 - \gamma) V_D(t'_{hh}; l) \leq V_D(t_l^{DM}; l). \quad (44)$$

4) C prefers DM' to DM assuming that the low-type D rejects DM' :

$$\begin{aligned} & \gamma V_C(t'_{hh}) + (1 - \gamma) V_C(t'_{hh}) \\ & > \gamma V_C(t_{hh}^{DM}) + (1 - \gamma) V_C(t_{hh}^{DM}). \end{aligned} \quad (45)$$

Suppose that DM' satisfies these conditions. Now consider a mechanism, DM'' , where $t_{ll} = t_{lh} = t_l^{DM}$, $t_{hl} = t'_{hl}$, and $t_{hh} = t'_{hh}$. This mechanism is incentive compatible (i.e., a low-type D reveals its type truthfully) due to (43). Moreover, conditions (42) and (45) imply that DM'' result in a higher ex ante expected joint welfare than DM . But this is contrary to the fact that DM is optimal.

Case 3: Only the low-type D and C prefer mechanism DM' . This case prevails iff the following four conditions are satisfied:

1) The low-type D prefers DM' to DM :

$$\gamma V_D(t'_{ll}; l) + (1 - \gamma) V_D(t'_{lh}; l) \geq V_D(t_l^{DM}; l). \quad (46)$$

2) The high-type D prefers DM to DM' :

$$\gamma V_D(t'_{hh}; h) + (1 - \gamma) V_D(t'_{hh}; h) \leq \gamma V_D(t_{hh}^{DM}; h) + (1 - \gamma) V_D(t_{hh}^{DM}; h).$$

3) C prefers DM' to DM given that the low-type D will also approve A :

$$\gamma V_C(t'_{ll}) + (1 - \gamma) V_C(t'_{lh}) \geq V_C(t_l^{DM}). \quad (47)$$

Lemma 5 implies that for any t'_{ll} and $t'_{lh} \in P_l$ there exists $t'_l \in P_l$ such that $\gamma V_D(t'_{ll}; l) + (1 - \gamma) V_D(t'_{lh}; l) = V_D(t'_l; l)$ and $\gamma V_C(t'_{ll}; l) + (1 - \gamma) V_C(t'_{lh}; l) > V_C(t_l^{DM}; l)$. But since $t_l^{DM} \in P_l$ there can exist no t'_l that is preferred to t_l^{DM} by both parties. Therefore, inequalities (46) and (47) cannot be satisfied simultaneously unless $t'_{lh} = t'_{ll} = t_l^{DM}$.

Therefore, since there is no alternative incentive compatible and renegotiation proof mechanism that is preferred at least weakly by both parties, DM will be impervious to interim renegotiations. ■

References

- Amador, M. and K. Bagwell (2013). The Theory of Optimal Delegation With an Application to Tariff Caps. *Econometrica* 81(4).
- Bagwell, K. (2008). Remedies in the WTO: An Economic Perspective. In V. Janow,

- M.E. Donaldson and A. Yanovich (Eds.), *The WTO: Governance, Dispute Settlement & Developing Countries*, pp. 733–770. Juris Publishing, Huntington, New York.
- Bagwell, K. and R. Staiger (1999). An Economic Theory of GATT. *American Economic Review* 89(1), 215–248.
- Bagwell, K. and R. Staiger (2005). Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause. *The Journal of Legal Studies* 34(2), 471–513.
- Baldwin, R. and F. Robert-Nicoud (2007). Entry and Asymmetric Lobbying: Why Governments Pick Losers. *Journal of the European Economic Association* 5(5), 1064–1093.
- Beaudry, P. and M. Poitevin (1995). Contract renegotiation: A simple framework and implications for organization theory. *Canadian Journal of Economics*, 302–335.
- Beshkar, M. (2010a). Optimal Remedies in International Trade Agreements. *European Economic Review* 54(3), 455 – 466.
- Beshkar, M. (2010b). Trade Skirmishes and Safeguards: A Theory of the WTO Dispute Settlement Process. *Journal of International Economics* 82(1), 35 – 48.
- Beshkar, M. and E. Bond (2012). Cap and Escape in Trade Agreements. *Working Paper*.
- Beshkar, M., E. Bond, and Y. Rho (2014). Tariff Binding and Overhang: Theory and Evidence. *Working Paper*.
- Bowen, T. (2011). Forbearance in optimal multilateral trade agreements. *Working Paper*.
- Brennan, J. R. and J. Watson (2013). The renegotiation-proofness principle and costly renegotiation. *Games* 4(3), 347–366.
- Buscaglia, E. and W. Ratliff (2000). *Law and economics in developing countries*. Hoover Institute Press.
- Buzard, K. and J. Watson (2012). Contract, renegotiation, and holdup: Results on the technology of trade and investment. *Theoretical Economics* 7(2), 283–322.
- Colares, J. (2009). A Theory of WTO Adjudication: From Empirical Analysis to Biased Rule Development. *Vanderbilt Journal of Transnational Law* 42(2), 383–439.

- Cooter, R. (1996). Decentralized law for a complex economy: the structural approach to adjudicating the new law merchant. *University of Pennsylvania Law Review* 144(5), 1643–1696.
- Dewatripont, M. (1989). Renegotiation and information revelation over time: the case of optimal labor contracts. *The Quarterly Journal of Economics*, 589–619.
- Goldsmith, J. and E. Posner (2005). *The limits of international law*. Oxford University Press, USA.
- Grossman, G. and A. Sykes (2007). United States - Definitive Safeguard Measures on Imports of Certain Steel Products. *World Trade Review* 6(1), 89.
- Hakobyan, S. (2013). GSP Expiration and Declining Exports from Developing Countries. *Working Paper*.
- Holmström, B. and R. Myerson (1983). Efficient and durable decision rules with incomplete information. *Econometrica: Journal of the Econometric Society*, 1799–1819.
- Klimenko, M., G. Ramey, and J. Watson (2007). Recurrent trade agreements and the value of external enforcement. *Journal of International Economics* 74(2), 475–499.
- Limão, N. and K. Saggi (2008). Tariff retaliation versus financial compensation in the enforcement of international trade agreements. *Journal of International Economics* 76(1), 48–60.
- Ludema, R. (2001). Optimal International Trade Agreements and Dispute Settlement Procedures. *European Journal of Political Economy* 17(2), 355–376.
- Maggi, G. (1999). The Role of Multilateral Institutions in International Trade Cooperation. *American Economic Review* 89, 190–214.
- Maggi, G. and R. Staiger (2011). The role of dispute settlement procedures in international trade agreements. *The Quarterly Journal of Economics* 126(1), 475–515.
- Maggi, G. and R. Staiger (2015). Optimal design of trade agreements in the presence of renegotiation. *AEJ: Microeconomics*.
- Maggi, G. and R. W. Staiger (2012). Trade disputes and settlement. *Working Paper*.
- Mattei, U. (1994). Efficiency in legal transplants: An essay in comparative law and economics. *International Review of Law and Economics* 3, 405–418.

- Neeman, Z. and G. Pavlov (2013). Ex post renegotiation-proof mechanism design. *Journal of Economic Theory* 148(2), 473–501.
- Park, J. (2011). Enforcing international trade agreements with imperfect private monitoring. *The Review of Economic Studies* 78(3), 1102–1134.
- Pelc, K. J. (2009). Seeking escape: The use of escape clauses in international trade agreements. *International Studies Quarterly* 53(2), 349–368.
- Reinhardt, E. (2001). Adjudication without Enforcement in GATT Disputes. *Journal of Conflict Resolution* 45(2), 174–195.
- Rosendorff, B. (2005). Stability and Rigidity: Politics and Design of the WTO’s Dispute Settlement Procedure. *American Political Science Review* 99(03), 389–400.
- Rosendorff, B. and H. Milner (2001). The optimal design of international trade institutions: Uncertainty and escape. *International Organization* 55(04), 829–857.
- Sykes, A. (1991). Protectionism as a Safeguard: A Positive Analysis of the GATT Escape Clause with Normative Speculations. *University of Chicago Law Review* 58(1), 255–305.
- Sykes, A. (2003). The Safeguards Mess: A Critique of WTO Jurisprudence. *World Trade Review* 2(03), 261–295.
- Sykes, A. (2006). *The WTO Agreement on Safeguards: A Commentary*. Oxford University Press, Oxford.
- Watson, J. (2007). Contract, mechanism design, and technological detail. *Econometrica* 75(1), 55–81.
- Wilson, B. (2007). Compliance by WTO members with adverse WTO dispute settlement rulings: the record to date. *Journal of International Economic Law* 10, 397–403.
- WTO. *Canada - Measures Affecting the Importation of Milk and the Exportation of Dairy Products - Notification of Mutually Agreed Solution*. Volume WT/DS103/33 and WT/DS113/33. World Trade Organization.
- WTO. *Dispute Settlement Understanding*. World Trade Organization.
- WTO. *European Communities - Measures Concerning Meat and Meat Products (Hormones)- Arbitration Report*. Volume WT/DS26/ARB. World Trade Organization.
- WTO (1995). The WTO Agreement on Safeguards.