Mahlstein, Kornel; Schropp, Simon A. B.

Working Paper

The optimal design of trade policy flexibility in the WTO


Provided in Cooperation with:
International Economics Section, The Graduate Institute of International and Development Studies

Suggested Citation: Mahlstein, Kornel; Schropp, Simon A. B. (2007) : The optimal design of trade policy flexibility in the WTO, HEI Working Paper, No. 27/2007, Graduate Institute of International Studies, Geneva

This Version is available at:
http://hdl.handle.net/10419/77437

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public. If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The Optimal Design of Trade Policy Flexibility in the WTO

Kornel Mahlstein
Graduate Institute of International Studies

Simon A.B. Schropp
Graduate Institute of International Studies

Abstract

This paper is a contribution to the literature on rational design of trade agreements. The World Trade Organization (WTO) is an incomplete contract among sovereign states. Incomplete contracts contain gaps. Ex post, contractual gaps may leave gains from trade unrealized; they may create “regret” in signatories once unanticipated contingencies or sudden protectionist backlashes have occurred. Trade policy flexibility mechanisms, such as the “safeguards clause” under Art. XIX GATT, are geared towards seizing ex post regret by allowing parties affected by a protectionist shock to partially and temporarily withdraw from previously made trade liberalization concessions – given that they compensate the victim(s) of such backtracking behavior.

This paper examines the somewhat understudied issue of optimal trade policy flexibility design in the WTO: In particular, we analyze whether ex post escape should be organized by means of a unilateral opt-out clause (a “liability rule” of escape), or a bilateral renegotiation provision (a “property rule” of escape). Modeling the WTO as a fully non-contingent tariff liberalization contract with contingencies (or “states of nature”) asymmetrically revealed, we find that a liability rule backed by expectation remedies payable to the affected victim Pareto-dominates both a renegotiation clause, as well as any other remedy arrangement connected to a liability rule. Only the remedial design of liability-cum-expectation damages yields the desirable incentives to liberalize ex ante, and to default ex post and therewith is able to replicate the outcomes of the hypothetical contracting ideal of the complete contingent contract.

© The Authors.
All rights reserved. No part of this paper may be reproduced without the permission of the authors.
The Optimal Design of Trade Policy Flexibility in the WTO

Kornel Mahlstein§ and Simon A.B. Schropp*

PRELIMINARY DRAFT (DECEMBER 07): PLEASE DO NOT QUOTE OR CITE

Abstract

This paper is a contribution to the literature on rational design of trade agreements. The World Trade Organization (WTO) is an incomplete contract among sovereign states. Incomplete contracts contain gaps. Ex post, contractual gaps may leave gains from trade unrealized; they may create “regret” in signatories once unanticipated contingencies or sudden protectionist backlashes have occurred. Trade policy flexibility mechanisms, such as the “safeguards clause” under Art. XIX GATT, are geared towards seizing ex post regret by allowing parties affected by a protectionist shock to partially and temporarily withdraw from previously made trade liberalization concessions – given that they compensate the victim(s) of such backtracking behavior.

This paper examines the somewhat understudied issue of optimal trade policy flexibility design in the WTO: In particular, we analyze whether ex post escape should be organized by means of a unilateral opt-out clause (a “liability rule” of escape), or a bilateral renegotiation provision (a “property rule” of escape). Modeling the WTO as a fully non-contingent tariff liberalization contract with contingencies (or “states of nature”) asymmetrically revealed, we find that a liability rule backed by expectation remedies payable to the affected victim Pareto-dominates both a renegotiation clause, as well as any other remedy arrangement connected to a liability rule. Only the remedial design of liability-cum-expectation damages yields the desirable incentives to liberalize ex ante, and to default ex post and therewith is able to replicate the outcomes of the hypothetical contracting ideal of the complete contingent contract.

Keywords: Incomplete contracts, remedies, enforcement, WTO

JEL classification: F02, F13, F51, F53, F55, K00, K33, K42

---

§ HEI Geneva; kornel.mahlstein@gmail.com
* University of St. Gallen (HSG), HEI Geneva, and NCCR Democracy; simon.schropp@gmail.com

Acknowledgements: The authors would like to thank Cédric Dupont, Heinz Hauser, Henrik Horn, Patrick Low, Petros Mavroidis, Damien Neven and Joost Pauwelyn for valuable discussions and vital input. All flaws, lapses, and errors remain the authors’. Schropp gratefully acknowledges financial contributions from NCCR-Democracy and the Graduate Institute, Geneva.
A. Introduction and research questions

It is an undisputed fact in the literature that the World Trade Organization (WTO) is a contract between sovereign nations.\(^1\)\(^2\) The insight that this contract is *incomplete* in many important aspects has recently gained acceptance among WTO scholars. There is a rapidly expanding literature that discusses and/or models the WTO as an incomplete trade accord between sovereign nations (recent contributions include Dunoff and Trachtman 1999; Ethier 2001; Kucik and Reinhardt 2007; Horn et al. 2006; Rosendorff 2005; Lawrence 2003; Bagwell 2007, Bagwell and Staiger 2005, Howse and Staiger 2005). A defining feature of incomplete contracts is that they contain gaps: Important contingencies (future conditions, or “states of nature”) are not considered in the terms of the original contract, and thus are not exhaustively and unambiguously specified *ex ante*, i.e. at the time of the conclusion of the contract. *Ex post*, during the performance phase of the contract, gaps may leave gains from trade unrealized. This, in return, may create room for “regret” (Goetz and Scott 1981) whenever unanticipated and unforeseen developments, or shocks, occur, such as a protectionist backlash within a country. A shock may seriously threaten some domestic import-competing sector or export industry, and therewith jeopardize welfare and/or employment of certain groups of society, or economic growth and social cohesion at large. Performance as previously agreed upon may then be neither desirable for the affected WTO Member nor mutually efficient anymore.\(^3\)

The fathers of the WTO were aware of the acute presence of contractual gaps and the inevitable uncertainty in the economic environment.\(^4\) In order to seize gains from *ex post* regret and to deflate the build-up of domestic pressure against trade liberalization, the WTO contract provides for certain trade policy flexibility instruments that permit one party (the “injurer”) to partially default, i.e. to step back, or withdraw, temporarily from contractual performance as previously agreed.\(^5\) The injurer can do so if

\(^1\) The WTO Appellate Body (AB) in *Japan – Alcoholic Beverages* expressly stated that “the WTO Agreement is a treaty – the international equivalent of a contract” (WT/DS 8,10,11/AB/R: 16, emphasis added).

\(^2\) As a convention, we will hitherto use the term *WTO* or the *Agreement* as shorthand for the bundle of multilateral contracts that are known as the Uruguay Round Agreements. These Agreements include the Marrakech Agreement (“WTO Agreement”, or “WTO Charter”), and all the accords in Annexes 1-4, such as the General Agreement on Tariffs and Trade (GATT), the General Agreement on Trade in Services (GATS), the Agreement on Trade-Related Intellectual Property Rights (TRIPS), the Agreement for the Settlement of Disputes (DSU), and others.

\(^3\) A signatory experiences regret whenever an *ex ante* envisioned transaction value is not realized in light of the newly revealed information. An unanticipated contingency arises which – had it been known to signatories *ex ante* – would have changed the content of the contract. Confer Mahoney (1999) at pp. 117:

> “A contract is an exchange of promises […] and the parties enter into it because each values the thing received more than the thing foregone. These values are based on expectations about the future because some or all of the contractual performance will occur in the future. When the future diverges from what a party expected, he may conclude that the performance he will receive under the contract is no longer more valuable than the performance he must provide. He has […] experienced a ‘regret contingency’ and now would prefer not to perform and not to receive the promised performance from the other party.”

\(^4\) The reasons why a multilateral trade agreement such as the WTO contains gaps shall not be discussed in detail. Generally, contracts are assumed to be incomplete for one of three reasons (or a combination thereof): (i) Asymmetrical information (Tirole 1994, Salanié 1997, chapter 7, Maskin and Tirole 1999); (ii) signatories’ rational cost-benefit calculations in connection with writing a complete contingent contract (Battigalli and Maggi 2002, Horn, Maggi and Staiger 2006); (iii) limitations of rationality in signatories, which manifest themselves in contracting errors, blunders and omissions (e.g. Schwartz 1992, Mahoney 1999, Masten 1999).

\(^5\) Trade policy flexibility tools are sometimes also called “trade contingency measures”, “safety valves”, or “escape clauses”. Later on it will become clear why neither of these terms is sufficient in covering the entire range of trade policy flexibility mechanisms.
certain (pre-)conditions are met, most notably that of compensating the parties affected from such back-tracking behavior (the “victims”). The WTO contains several formal, de iure, trade policy flexibility tools, such as Art. XIX of the GATT, which regulates the use of safeguards. In addition to these de iure escape clauses there are various informal, de facto, flexibility tools available to WTO Members. Examples are countervailing duties or antidumping action against allegedly unfair trade of other WTO Members.

There is an abundant literature on the WTO’s formal and informal escape mechanisms (see Schropp 2008, chapter 1, for a detailed literature review). One elemental question, however, remains understudied: What is the optimal design of a remedial regime in a multilateral trading agreement such as the WTO? How should flexibility mechanisms best be organized? Intuitively, from the outset of negotiations, signatories will want to ensure (politically) optimal ex ante trade liberalization commitments. Yet at the same time, contracting parties want to guarantee optimal flexibility, once unforeseen contingencies render rigid compliance with the original contract inefficient. At the same time, ex post flexibility should not be open to opportunistic abuse as a protectionist tool. Heeding these directives, thus, is it prudent for WTO Members to allow for ex post escape at all, or would not signatories rather be better off with a mandatory contract performance obligation? Is trade policy flexibility in the WTO best organized as a ready-to-use escape mechanism that allows any injurer party to (partially) “opt out” of his contractual performance obligation at any point in time – or is it more prudent to structure trade policy flexibility as ex post renegotiations between injurer and victim (“buy-out”)? Furthermore, which contractual compensation remedies are to be awarded to the victim of such a backtracking measure? Should the victim be put in as good a position as if the injurer had performed? Should the status quo ante the breach be re-established, or rather the status quo ante the contract? Alternatively, should the victim rather receive a fair share of the actual efficiency gains generated from the injurer’s temporary withdrawal from his contractual obligations?

In order to find convincing answers to these questions, we develop a model of the WTO as an incomplete tariff liberalization contract where future states of nature are asymmetrically revealed to the signatories. Modeled as a cooperative (enforceable) game, we endogenize the choice of trade policy flexibility regime. Negotiating behind a Rawlsian veil of ignorance, countries agree upon a specific trade flexibility rule (comprising of a rule of non-performance and the accompanying remedy), before starting an infinitely repeated trade-setting game in an uncertain world. We compare the differential effect of various escape mechanisms on trade liberalization (ex ante), and countries’ decision to engage in contractual breach (ex post). More specifically, we evaluate the incentive-compatibility of (i) mandatory compliance (prohibition of ex post escape), (ii) a rule of renegotiation, (iii) expectation damages (compensation for efficiency losses), (iv) reliance damages (re-establishment of the status quo ante the breach), (v) restitution (re-establishment of the status quo ante the contract), and (vi) no compensation.

Two aspects make this research endeavor worthwhile: First, we believe that the current system of trade policy flexibility in the WTO is profoundly flawed, and in need of substantial reform. Second,

---

6 We define victim and injurer as roles, not as conditions. This is to say that any contracting party experiencing some form of contractual regret or doubt assumes the role of the injurer independently of whether it acts upon that regret (e.g. by partially withdrawing from its obligations) or not. Equivalently, the role of the victim is that of the actor affected by the other party’s regret or doubt. Both terms are non-judgmental and are free of welfare implications to the players (a victim is not necessarily worse off by assuming that role). For the ease of reading, and to avoid misunderstandings the injurer throughout this paper will be referred to as he, while the victim carries the personal pronoun she.
contrary to the fields of law & economics (L&E) of commercial law, or industrial organization (IO), WTO scholarship so far has failed to adequately address the issue of optimal design of breach and remedy in an incomplete trade contract.

The remainder of the paper is organized as follows: Part B will give a general introduction to flexibility in incomplete contracts, and trade agreements in special. This is followed by an assessment of the current system of trade policy flexibility in the WTO. We argue that the reality of trade policy flexibility in the WTO defies much of what contract theory would positively prescribe. We conclude that the current WTO design of flexibility is in need of repair. Part C provides a short literature review: Models of breach and remedy in WTO scholarship and neighboring fields are compared to our approach. Part D forms the core of the paper: We introduce a model of the WTO as a fully non-contingent tariff liberalization treaty, in which contingencies are asymmetrically revealed. Modeled as a cooperative on-the-equilibrium game, we compare different trade policy flexibility regimes in the presence of real-life transaction costs. Part E concludes and gives policy recommendations.
B. Flexibility in trade contracts

In the first part of this section, we delineate the general concept of flexibility in incomplete contracts, and then proceed to apply those insights to multilateral trade agreements. Finally, we give a structured critique of the current trade policy flexibility regime of the WTO.

B.1 The essence of flexibility in incomplete contracts

In a seminal contribution to L&E, Calabresi and Melamed (1972) articulated that signatories to a contract are “making their own law” by exchanging residual ownership rights, or “entitlements”. Contracting parties mold mutual commitments into a set of reciprocal rights and obligations that capture – as unambiguously as possible – nature, extent, and limits of the agreed-upon cooperation.

Parties thereby generally perform three distinct steps (cf. also Pauwelyn 2006, Trachtman 2006): First, by crafting the “primary rules” of contracting, signatories define and assign the mutual entitlements. Primary rules are the essence of the contract; they establish the “depth” and “breadth” of cooperation.

In a second stage, contracting parties agree on rules of entitlement protection: These “secondary rules of contracting” allocate residual decision rights which organize ex post behavior during the contract performance phase. After having delineated intra- contractual, permissible, non-performance behavior, transactors in a third step accord how to sanction extra- contractual, uncooperative behavior. These “tertiary rules” are the enforcement mechanisms and procedures taken in response to the illegitimate taking or destruction of an entitlement.

Secondary rules of entitlement protection lay out the scope for ex post trade flexibility, since they define the legal limits of escape. Intuitively, the more incomplete a contract, the more important is...
the careful design of secondary rules.\textsuperscript{11} Every contractual flexibility mechanism consists of a rule of “breach”, as well as of an accompanying remedy rule.\textsuperscript{12}

Three generic types of contractual flexibility are usually described in the literature: (i) inalienability-, (ii) liability- and (iii) property rules of entitlement protection:

(i) Whenever an \textit{ex post} entitlement transfer (to take, sell, or trade residual rights) is considered inefficient or immoral, a rule of \textit{inalienability} (also termed “mandatory specific performance”) is agreed upon by signatories. A rule of inalienability (henceforth IR) mandates unconditional specific performance of the contractual rules, no matter what contingencies may occur in the course of contract performance. The protection of the initial entitlement allocation is then called absolute, and a taking of the entitlement is strictly prohibited.\textsuperscript{13}

(ii) If contracting parties generally consent to the possibility of trading or reallocating entitlements \textit{ex post}, they must determine whether to protect initial entitlements by means of a \textit{liability rule} (LR) or a \textit{property rule} (PR) of entitlement protection. Under a pure LR, one party (the taker or injurer) has the option to unilaterally take away parts of the other party’s entitlement without the latter’s prior assent. The taker can engage in the unilateral appropriation (which effectively is an expropriation of the holder of the entitlement) under the condition that he compensates the owner (or victim) for damages suffered – usually by paying a previously specified exercise price.

(iii) Under a property rule of protection both parties are under a strict obligation to respect the initial entitlement distribution. The injurer is directed to perform, and a failure to do so will be punished severely. However, a party experiencing regret can buy off the owner’s entitlement through \textit{renegotiations}. The injurer (rather: the requesting party) can avoid his commitments by securing permission from the victim (rather: concerned party or owner) – usually by paying a bargained amount. Whenever the parties come to an agreement, the owner cedes his entitlement and sells it to the requesting party – the transfer is thus \textit{bilateral}.

Any rule of breach must in all cases be accompanied by a corresponding remedy rule. At the conclusion of the contract signatories agree on how “costly” \textit{intra-contractual} non-performance by the injurer (read: breach) should be. The economic function of remedies in contracts clearly is to alter the incentives of the regret party and to grant the victim indemnity for the other party’s (partial) default. There is a continuum of remedies ranging from a zero-damage payment to coercive (infinitely high) damages (Mahoney 1999, pp.121). The following damage remedies warrant brief explanation:

\begin{itemize}
  \item [\textsuperscript{11}] If – hypothetically – a contract were complete in that it specified in detail \textit{all} possible contingencies and prescribed comprehensive plans of actions, flexibility mechanisms would be superfluous. Every \textit{ex post} non-performance then would by definition be \textit{extra-contractual}, i.e. deviating, punishable behavior.
  \item [\textsuperscript{12}] The term \textit{breach} is somewhat misleading, since in everyday terminology \textit{contract breach} bears the connotation of \textit{extra-contractual}, illegal behavior. However, in contract theory, \textit{breach} is often used to describe \textit{lawful} escape. As a convention we will use \textit{breach} as in the latter (cooperative) sense, and use the term \textit{non-performance} as a generic term for any kind of (\textit{intra-or extra-contractual}) \textit{ex post} default.
  \item [\textsuperscript{13}] Political rights are an example of inalienable rights that cannot be traded: It is not legally possible to sell off or trade your vote in a ballot.
\end{itemize}
• The restitution remedy re-establishes the status quo ante the contract. An injurer must restore the Nash-level that persisted before the contract in the non-cooperative past.\textsuperscript{14} In other words, restitution damages aim at reconstructing what the contract would look like if the measure in question had never been part of the initial deal.

• The reliance remedy obliges the injurer to re-establish the victim’s status-quo ante the breach.\textsuperscript{15} It aims to reestablish a world that would exist if the illegality were removed. One conceivable way of “paying” reliance damages is to seize a contested measure and return to fully cooperative behavior.

• The expectation remedy places the victim in as good a position as it would have been had the injurer performed his contractual obligations. It is equivalent to the replacement value that exactly makes the victim indifferent between the injurer’s performance and his default.\textsuperscript{16} Expectation damages are essentially what is called “commensurate punishment” in game theory: They insure the victim against any dynamics that unfold \textit{ex post}.\textsuperscript{17}

• Efficiency damages, next, are expectation damages plus the efficiency gains from non-performance. Efficiency damages reconstruct a world that would exist if all the gains from having enacted the measure in question were given to the victim. This actually puts the victim in a better position than it had been, had the injurer performed as promised.

• Negotiated remedies are agreed upon after the occurrence of a contingency pursuant a liability rule of flexibility. In principle, negotiated damages can lie anywhere between zero and coercive remedies.\textsuperscript{18}

Note that not all combinations of escape and remedy rules are equally meaningful: Inalienability rules are best adhered to by coercive penalties (prison, forfeit, coercive liquidated damages, etc.). Property-rule protection logically is accompanied by bilaterally negotiated remedies in excess of expectation damages.\textsuperscript{19} Liability-rule protection can be accompanied by various forms of damage measures.

\textsuperscript{14} The pre-contractual Nash-level may or may not correspond to the zero-cooperation level. Oftentimes restitution damages mandate the injurer to pay back any upfront reliance investments made by the victim in anticipation of the contract (e.g. the construction of additional production facilities).

\textsuperscript{15} Reliance is presumably a larger amount for the victim than restitution damages, since prior transaction efficiencies have to be accounted for. The longer the contractual relationship has been going on, the more restitution and reliance measures differ.

\textsuperscript{16} In other words, expectation damages aim to reconstruct what the world would look like if the measure in question had never been committed.

\textsuperscript{17} The difference between expectation and reliance damages is that reliance compensates the victim for direct harm suffered, but leaves aside indirect effects and foregone opportunities. Mavroidis (2000 at p. 800) argues that expectation damages are strictly higher than \textit{damnum emergens} (direct harm suffered). Rather, they must be interpreted as \textit{lucrum cessans}: \textit{All} further efficiency costs (opportunity costs or losses in value-added) caused by the partial breach of the agreement over and above direct effects must be indemnified. Indeed, it is exactly these contractual efficiency gains that motivated transactors to conclude the contract in the first place (see also Schropp 2005 at footnote 10 and accompanying text).

\textsuperscript{18} Signatories also often opt for the liquidated damages measure, which is a contractually fixed (\textit{ex ante} negotiated) indemnity. A liquidated damages clause usually refers to a certain sum or action which falls due whenever one signatory deviates from its promised behavior (Masten 1999, p. 27). While liquidated damages are an interesting concept, it is difficult to generalize their impact and outcome, simply because their design and magnitude are unique to each contractual setting.

\textsuperscript{19} Under a PR, negotiated damages can be expected to lie between the expectation and the efficiency measures: The rational victim will not settle at a loss (and expectation damages put her into a position where she is exactly indifferent between performance and non-performance), and ideally would want to capture all the injurer’s efficiency gains from non-performance for her willingness to “let go” of its contractual rights.
Practically every contract in existence is incomplete in at least some important aspects. Prior to the conclusion of a contract, no contracting party possesses full knowledge of nature and impact of future contingencies, nor of the possible policies and instruments that its partners might concoct during the entire life span of a contract. Also, it is difficult to anticipate whether future contingencies will make a signatory victim or injurer. A complete contingent contract (henceforth CCC) must remain an unachievable ideal. In the absence of post-contractual transaction costs (litigation-, renegotiation-, enforcement- and policing costs) and asymmetrical revelation of information, the choice of flexibility regime is inconsequential. In such a world any initial allocation leads to an efficient outcome through ex post renegotiation. However, the design of an ex post non-performance regime tremendously gains in importance, once conditions are less clinical and real-life imperfections are taken into consideration.

The benchmark of flexibility in any incomplete contract is determined by the achievable first-best, which we shall call “efficient breach contract” (EBC). The EBC is the one arrangement that mimics the outcome (but not the substance!) of the hypothetical complete contingent contact (cf. Schropp 2008, section 3.4). In doing so, the EBC must fulfill two critical properties: (i) that of efficient non-performance (or efficient breach) ex post, and (ii) that of efficient cooperative commitment ex ante.

(i) The efficient breach property safeguards that signatories use a flexibility instrument if, and only if, it is ex post efficient to do so. Efficient non-performance ensures the optimal balance between flexibility and entitlement protection: Every contractual entitlement is protected in a way that prohibits opportunistic opt-out, yet reaps the opportunities opening up as a consequence of the occurrence of unforeseen regret contingencies. This is equivalent to saying that injurers will only engage in flexibility so as to seize welfare-enhancing ex post non-performance opportunities – without harming the victims of contractual escape.

(ii) In addition, any efficient flexibility instrument must ensure the optimal scale and scope of ex ante cooperation. After all, cooperative commitments determine the gains to be had from the contract in the first place. They are thus the currency of any agreement. If signatories grant each other concessions that only marginally differ from the non-cooperative Nash-game, the contractual gains in consequence are also minimal.

---

20 The contracting ideal for every contract is the Pareto-optimal complete contingent contract (CCC). A CCC is the first-best contract between signatories (see Shavell 1980). It is an Arrow-Debreu-type accord that completely informed, perfectly rational parties would write in the absence of any contracting imperfection, such as negotiation costs, costs of information gathering, asymmetrical knowledge, or bounded rationality (e.g. Masten 1999, p. 27). It provides a comprehensive description of every possible present and future state of the world. A CCC hence spells out in complete detail the exact legal rights and duties of each party – no matter how small the probability of the contingency (Cohen 1999, p. 79). It exhaustively prescribes actors’ efficient actions (Hart and Moore 1988, Posner 1988, Craswell 1999, Masten 1999) and is therefore apt to exploit all possible gains from trade..

21 This, of course, is an application of the Coase Theorem (cf. Coase 1937, Kaplow and Shavell 1996b, Dunoff and Trachtman 1999, p. 23).

22 In contract theory the nomenclature “efficient breach” is sometimes used for the flexibility design of liability rule-cum-expected damages (e.g. Friedman 1989). We refrain from this usage, because we believe it to be misleading: Efficient adjustment to changing circumstances and unforeseen contingencies is in principle possible under each of three flexibility arrangements of liability, property or inalienability rule. Confer Sykes (2000, p. 353): “To be sure, it is possible for efficient breach to occur with other remedial options [than the LR]. In particular, if specific performance is the remedy for breach a party wishing to breach can always approach the other party and attempt to negotiate a release of performance”.

— 7 —
Whenever one of these two conditions fails to hold, the contractual system of non-performance is out of balance, the objective of the EBC is missed, and the expected value of the contract is likely to be suboptimal. Intuitively, if existing flexibility provisions are unsuccessful in safeguarding the initial commitment level, a signatory suspecting to assume the role of a victim is likely to commit to less cooperation upfront. On the other hand, the prospect of rigid flexibility clauses or overzealous (punitive) compensation, and consequently of un-seized instances of regret, discourages future injurers from cooperating. At the same time, it may prompt future victims to over-commit, that is, to invest an inefficient amount of resources in the contract.

**B.2 Trade policy flexibility in incomplete trade contracts**

The WTO, like any trade agreement, is a contract between sovereign nations. Those who hold political office – incumbent trade policymakers – sit down and negotiate trade liberalization deals. The primary rationale for contracting is the *reciprocal exchange of market access*:23 Trade policymakers, like any rational actor, are self-interested utility maximizers, or *homines oeconomici* (cf. Kirchgässner 2000). Although these individuals are not necessarily happy granting trade partners additional access to their home markets, they are nevertheless better off by signing reciprocal trade liberalization agreements that safeguard competitive opportunities and generally less distorted market access for their own exporting industries.

In the WTO the market access that countries are willing to grant to each other is laid down in and constituted by the compulsory tariff bindings (the “schedules of commitments” regulated in Arts. II and XXVIII*bis* GATT), and by positive GATS concessions in the four service modes. In addition to these reciprocal trade liberalization concessions countries lay down specific modalities aimed at preserving the mutually agreed balance of concessions, and at prohibiting the use of trade policy instruments apt to partially undo this initial equilibrium of rights and obligations. Examples for these negative provisions are non-discrimination stipulations (as in Art. I and III GATT), the prohibition of quantitative restrictions (as in Art. XI GATT), codes of conduct associated with non-tariff barriers and red-tape measures (e.g. Art. III GATT, SCM, TBT, SPS, GPA, ROO24), and explicit exceptions to the right of competing in foreign markets.25

These explicit provisions, however, cannot prevent the WTO, or any other trade agreement for that matter, from being a highly incomplete contract (Schropp 2008, chapter 4.3). A trade policymaker cannot predict with certainty (or time is too valuable) future states of the world and how they will affect the country she is representing. She cannot foresee which country will be affected by what kind of environmental shock – be it price or supply shifts, an exchange rate shock, a production technology innovation, or an endogenous change in a country’s political institutions or its constellation of political

---


24 The acronyms stand for the “Agreement on Subsidies and Countervailing Measures”, the “Agreement on Technical Barriers to Trade”, the “Agreement on Sanitary and Phytosanitary Measures”, the “Agreement on Government Procurement”, and the “Agreement on Rules of Origin”, respectively.
majorities. More serious even, an exogenous and unanticipated shock may affect domestic firms’ demand, or ability to lobby, for protection of their markets. Thus, no trade policymaker can predict with accuracy the complex internal reverberations that any external shock will bear on her/his political support function.

Unforeseen and unforeseeable contingencies may make it impossible for a signatory country, respectively its government, to honor its previously made trade liberalization commitments in all sectors at all future points in time. So a key question for signatories of a trade agreement effectively is: Which is the most efficient governance structure of non-performance? In other words, which institutional design arrangement can concomitantly achieve sufficient flexibility and an efficient level of upfront commitments and thereby match the requirements of the ideal EBC? An efficient system of trade policy flexibility must prevent (i) opportunism by the injurer, (ii) un-seized regret contingencies on the part of the injurer, and (iii) insufficient compensation paid to the victim. Any one of these potential pitfalls can be expected to have serious repercussions on the 

B.3 Trade policy flexibility in the WTO

Current practice of trade policy flexibility in the WTO is in need of improvement. We can only sketch our critique of the existing WTO flexibility regime here.

The WTO is endowed with several formal, de iure, trade policy flexibility mechanisms that permit ex post modification of previously made trade concessions. The GATT provides for so-called “contingency measures”, like Art. XII (Restrictions to Safeguard the Balance of Payments, applicable to developed countries), Art. XVIII (infant industry protection and balance of payments crises; applicable to developing countries), Art. XIX (Emergency Actions on Imports of Particular Products, also known as “safeguards clause”), Art. XX (General Exceptions), or Art. XXI (Security Exceptions). GATT Art. XXVIII (Modification of Schedules, also known as “tariff renegotiation”) is the default or

25 Examples for exceptions include GATT Arts. IV (on cinematography), XVII (on state trading), Art. XXIV GATT (on preferential trade agreements), XXI (on national security), the “Enabling Clause” (based on GATT Arts. XXXVI-XXXVIII), and waivers (Art. IX of the WTO Agreement).

26 Opportunism has been defined as “self-interest seeking with a guile” by Williamson (1979, p. 234, 1985 at p. 47). A more precise definition to our mind is “inefficient redistribution” (Cohen 1999, p. 90): In contrast to a lump-sum transfer, opportunism is always welfare-depreciating at the expense of a victim party.

27 In other words, any efficient trade policy flexibility regime must overcome what Ethier (2001, p. 5) calls the reciprocal-conflict problem, which he describes as follows:

“Each country is aware, ex ante, that it may find itself, ex post, harmed by a policy that some trading partner wishes to make. So the former will want a recognized punishment procedure as a deterrent. But that country will also be aware, ex ante, that it might find, ex post, itself in a position where it would be costly not to take some policy action that would harm a partner. This is the reciprocal-conflict problem: Every country knows that it might turn out to be either the accuser or the accused. Thus it is in no country’s best interest, ex ante, to agree that, ex post, either the accuser should be unconstrained in its ability to punish or the accused should be unconstrained in its ability to proceed without punishment” (emphases in original).

28 The discussion of trade policy flexibility mechanisms in this subsection shall be reduced to trade in goods covered by the GATT (and appending Agreements). Similar instruments for the trade in services are provided for in the GATS.

29 The interested reader is referred to more comprehensive reviews of the GATT/WTO system of non-performance, such as Schropp (2008, section 5), Roitinger (2004); Kleen (1989), or Finger (1998).
fallback rule of flexibility. It can be invoked independently of any contractually anticipated contingency; political expediency by the injurer (the “requesting Member”) suffices to kick off tariff renegotiations at any time.\(^{30}\)

We have several issues with the WTO’s formal flexibility regime:

1. **Article XXVIII GATT is a dubious fallback rule of flexibility.** First, the scope of Art. XXVIII is limited to permanent, not temporary, tariff modifications. The procedural build-up connected to tariff renegotiations renders emergency and temporary protection impossible.\(^{31}\) Second, the renegotiation provision of Art. XXVIII GATT is foiled by a weak remedy: Contract theory would prescribe a rule of renegotiation to be protected by high enforcement remedies in order to be operable. Anything less reduces a property rule to a liability rule, because the injurer will practically always choose to be punished instead of engaging in renegotiations. Yet, this is exactly the case with Art. XXVIII: Although at first sight the provision comes across as a property rule of flexibility backed by negotiated compensation (paras. 1 and 2), Art. XXVIII is effectively a liability rule-type opt-out accompanied by WTO arbitrators’ interpretation of commensurate damages.\(^{32}\)

2. **No substantial difference between GATT Arts. XXVIII and XIX.** The finding that tariff modification is actually a liability-type escape provision then begs the question if there is a substantial difference between GATT Art. XXVIII and the whole arsenal of liability opt-outs, the most important one arguably being Art. XIX GATT. Comparing Arts. XXVIII and XIX GATT, it is somewhat elusive why Members’ market access is protected by two flexibility mechanisms, both of which feature a liability rule and the same remedy (“substantially equivalent” damages, according to Arts. XIX.3 and XXVIII.2(a)).

3. **Insufficient scope of de iure escape clauses.** Many scholars are concerned by the fact that de facto violations of WTO obligations frequently occur because of the rigidity connected to the enactment of

---

\(^{30}\) Although it is nowhere stated *expressis verbis* that GATT Art. XXVIII is the default flexibility rule, we infer that this is actually the case: Whenever no other contingency measure applies, a country wishing to alter its previously negotiated market access balance is left with the possibility of tariff renegotiations as the sole flexibility solution.

\(^{31}\) If the renegotiation request of the requesting party does not happen to fall into an official triennial renegotiation period of three months (laid out in Art. XXVIII.1), the injuring Member has to secure the prior authorization of all Members in order to enact its right to renegotiate (Art. XXVIII.4). As per *Interpretative Note ad Art. XXVIII GATT* (§§4.1, 4.4) the requesting Member must submit a written request to the Council of Trade in Goods, the relevant organ to decide. The request must contain comprehensive statistical and other information justifying the Member’s appeal and a list of the effects of the envisaged measure. The Council will then give notice of its consensus decision within 30 days. *Interpretative Note ad Art. XXVIII GATT* at §4.5 states that later on in the process the same Council determines (again by unanimous consensus) whether the compensation offered by the injurer is sufficient. This sort of conditionality is certainly apt to slow down the process of reacting promptly to unforeseen contingencies and unanticipated shocks.

\(^{32}\) Paragraph 3 of Art. XXVIII states that an injuring Member may proceed to unilaterally withdraw concessions in cases where negotiations over MFN compensation break down. Adversely affected trading partners may then bilaterally retaliate by withdrawing substantially equivalent concessions or other obligations through recourse to binding arbitration (as per Art. 22.6 DSU). Pertinent parts of Art. XXVIII.3 read (emphases added): “If agreement between the Members primarily concerned cannot be reached […], the Member which proposes to modify or withdraw the concession shall, nevertheless, be free to do so and if such action is taken, any Member with which such concession was initially negotiated, any Member […having] a principal supplying interest and any Member […having] a substantial interest shall then be free not later than six months after such action is taken, to withdraw […] substantially equivalent concessions initially negotiated with the applicant contracting party.” This wording is little else than putting into effect a liability-type escape possibility for injurers. It arguably renders the previous renegotiation clause futile.
formal escape mechanisms, such as Art. XIX GATT. The current WTO safeguards regime does not provide Members with the necessary “breathing space” (see Horn and Mavroidis 2003, Sykes 2003, Roitinger 2004). As became clear in due course of the EC – *Hormones* cases, the European Communities – for political or health reasons – wished to step back from previously made commitments in the beef sector. This endeavor, however, is not considered in any formal WTO escape clause. Lacking any official means of withdrawing from existing concessions, the EC saw no alternative to keeping up the violation of the Agreement. Lacking any formal means to spontaneously and temporarily withdraw from existing concessions is a serious issue, because it effectively blurs the line between good-faith and bad-faith (opportunistic) behavior.

4. Retaliation is a questionable mechanism of remediation. Most *de iure* flexibility mechanisms couple *ex post* escape with the remedy instrument of tariff compensation offered by the injurer to the victim(s) of a backtracking measure. However, if compensation negotiations break down, the victim is authorized to engage in suspension of concessions or other obligations, also known as “retaliation” or “trade sanctions”. As is well known, the countermeasure of retaliation is a questionable, and some would say nonsensical, mechanism of remediation.

5. Arbitrators’ calculation of remedies is systematically under-compensatory. As per Arts. XVIII.7.b, XIX.3.a, or XXVIII.3.a/4.d, the WTO mandates *intra*-contractual remedies to be “substantially equivalent” to the damage done. Pursuant to Art. 22.6 DSU, it devolves upon WTO arbitration panels to interpret what exactly substantial equivalence means: By granting *prospective* remedies amounting to *direct trade damages*, WTO arbitrators have interpreted this term to roughly imply the re-establishment of the *status quo ante* the breach (Mavroidis 2000, Lawrence 2003, Breuss 2004, Spamann 2006): In constructing a counterfactual situation that would prevail if the illegality were removed (Sebastian 2007, p. 351), WTO arbitrators have awarded damages that are apt to restore the trade level that would exist, had the injurer brought his contravening measure into conformity after the “reasonable period of time”. Hence, the current system of contractual remedies is probably best characterized as a remedy *short of* reliance damage measure. As we will argue below in more detail, we think that arbitration panels have shown a somewhat defective understanding of commensurate damages. In a nutshell, our formal model will show that reliance remedies are under-compensatory. This has a bearing on victims’ willingness to liberalize, be it *ex ante* or in ongoing trade rounds.

33 “*Measures Affecting Livestock and Meat (Hormones)”* (DS 26 and 48), and “*Continued Suspension of Obligations in the EC — Hormones Dispute*” (DS 320 and DS 321).

34 This is of course but one way of interpretation. Other observers may come to different conclusions, arguing that the EC acted genuinely malevolently, or was “putting to a test” the infant dispute settlement system of the WTO.

35 Charnovitz (2001) lists at least 8 serious disadvantages of the countermeasure of retaliation. Most compellingly, a victim country “shoots itself in the foot” (Mavroidis 2000) when avenging the injurer’s unlawful trade protection by raising trade barriers on its part, since doing so is inefficient, harms consumers and downstream industries, and in addition makes import-competing industries complacent and slack.

36 We say “roughly” and “short of” the reliance damage measure, because the *prospective* nature of arbitration award renders impossible the exact re-establishment of the *status quo* of the market access balance as it existed before the breach. The remedy applied by WTO arbitrators fails to compensate the victim for direct trade damages incurred between the enactment of the measure in question and the lapse of the RPT (cf. for example Pauwelyn 2000).
In addition to de iure escape clauses just described, there are various de facto flexibility tools available to WTO Members. Things get more irritating, once these informal flexibility tools enter into consideration: Trade policy tools such as voluntary export restrictions (VERs), ordinary marketing agreements (OMAs), antidumping (AD) and countervailing duty (CvD) measures, subsidies, non-discriminatory domestic policies, or quite simply a violation of the Agreement, may be utilized by WTO Members as means of escaping initially made trade liberalization commitments and of partially undoing the original trade deal.\footnote{VERs and OMAs are easy to enact and relatively cheap for the injuring country. Their use was common in the 1970s and 80s (Bhagwati 1988, Baldwin 1989), but got prohibited in the course of the UR negotiations (by virtue of Art. 11.2 of the Safeguards Agreement, SGA). There are, however, good reasons to believe that VERs/OMAs are still in practice today: First, the proverb “no plaintiff, no judge” applies. How can an outsider find out whether exporters reduce their output due to, say, bottlenecks in production or due to a secret bilateral agreement? Rosendorff (1996), for example, points out that even today many VERs are the outcomes of antidumping investigations. Bown (2002b, p. 53) suggests that AD actions themselves offer a loophole for signatories to engage in managed trade and VER-like price undertakings. Second, VERs are legally possible through a loophole created by the application of Art. XIX GATT in connection with the footnote to Art. 11.2 and Art. 5.2 SGA (see Mavroidis 2007, chapter 4.7.5): Art. 5.2 SGA allows for discriminatory application of safeguards measures, whereas the footnote to Art. 11.2 SGA permits import quotas as safeguard measures to be administered by the exporter. This is little else than having VERs enter through the backdoor.} Given that they are in contravention of the letter of the law, or at least of the spirit of the Agreement, these de facto escape mechanisms (especially AD and CvD action, and violation of the Agreement) supersede de iure trade policy flexibility tools, simply because they are (1) easier to enact, (2) possess a broader scope of application, (3) are cheaper in terms of political currency, and (4) are inexpensive in terms of remedies payable to the victim.

A. Informal flexibility instruments have lower enactment thresholds. While WTO Members experiencing regret may invoke, say, GATT Art. XIX on safeguards only if a string of four rather rigid conditions are cumulatively met,\footnote{A country that wants to invoke Art. XIX GATT must show that “i) as a result of unforeseen development; ii) imports in increased quantities; iii) have caused or threatened to cause; iv) serious injury to the domestic industry producing the v) like product” (Howse and Mavroidis 2003 at p. 686, emphasis in original). For specifics of this five-prong legal test and additional restrictions for the enactment of Art. XIX by case law, see also Mavroidis (2007, chapter 4.7.4), or Roitinger (2004 at pp. 102). For example, according to the AB ruling of US – Wheat Gluten (WT/DS 166/AB/R at § 55), the injurer, when claiming “serious injury”, must take into consideration all relevant factors that determine injury.} informal flexibility instruments have considerably lower enactment thresholds – precisely because they are enacted in the shadow of the law: AD and CvD, for example, were written to counter unfair trading, not to be abused as tools of protectionism.\footnote{It is no secret that AD and CvD measures are quite frequently abused as protectionist tools of flexibility. Officially, those two tools can be invoked only to counter “unfair” trade practices of foreign exporters, subject to the occurrence of material injury caused by the allegedly unfair imports. Unofficially, these two measures are highly fungible mechanisms which are easily turned into protectionist instruments. The protectionist abuse of these trade remedy measures is the dominant perception in the trade literature (examples include Finger et al. 1982, Sykes 1989, Palmeier 1991a; 1991b, Schuknecht 1992, Tharakan and Waelbroeck 1994, Tharakan 1995, Messerlin 2000, Bown 2001, Finger et al. 2001, Neufeld 2001, Bown 2002b, Prusa and Skeath 2002, Barfield 2005, Trebilcock and Howse 2006).} Also, violation of the Agreement and the enactment of trade-related domestic policy measures can be utilized by injurers anytime and without any precondition.

B. Informal flexibility mechanisms possess a broader scope of application. De facto escape tools are less impeded by confining contractual language concerning their application scope. Market-closing domestic policy instruments, for example, can be enacted for all sorts of reasons, not just for balance-of-payment crises (as prescribed by Art. XII GATT). An important factor also is the issue of selectivity: Many informal escape mechanisms, most notably VERs, AD, CvD, and violation of the...
Agreement, can be pinpointed against those exporting countries, industries, or firms that are of special concern to an importing WTO Member. Instead of having to treat all countries in the same way, injurers can target their protectionist actions more precisely.\textsuperscript{40}

\section*{C. Informal trade policy flexibility tools are politically more convenient to policymakers.} Several authors have examined the political economy of flexibility mechanisms. They find various reasons why informal escape mechanisms (in particular AD, CvD, and violation of the Agreement) are politically more opportune to protectionist policymakers than formal flexibility tools.\textsuperscript{41} There is an intuitive explanation for the reluctance of countries to utilize \textit{de iure} escapes: First, policymakers are averse to overtly legitimizing a protectionist measure applied against fairly traded goods – this would be tantamount to an official admission of guilt and incompetence. Informal opt-outs such as AD/CvD action or VERs circumvent problems of this kind. Second, the \textit{direct} nature of the \textit{de iure} measures does not allow for “blame-shifting” or “scapegoating” of allegedly unfair foreign trade practices (Ethier 2004a, Schropp 2005). Third, formal flexibility tools display a lack of selectivity (confer footnote 40 above). Finally, the requirement that compensation be offered immediately – not after a lengthy litigation – militates against engaging in formal escape.

\section*{D. \textit{De facto} flexibility tools come with lower remedies.} It is important to appreciate that in the WTO \textit{intra-} and \textit{extra-}contractual behavior get sanctioned in the same way.\textsuperscript{42} We are hard-pressed to think of any logical reason why the framers of the GATT/WTO could have felt that this was a good idea.\textsuperscript{43} On the contrary, contract theory postulates that illegal \textit{ex post} behavior should be punished by high \textit{extra-}contractual remedies. By sanctioning \textit{extra-}contractual defection not tougher than contract-conform default, the WTO effectively reduces the distinction between lawful and prohibited behavior to a legalistic formality. What’s worse, deliberately violating the Agreement actually is penalized \textit{less} than if a Member resorts to a \textit{de jure} flexibility mechanism: Depending on which \textit{de facto} flexibility tool is chosen, applied WTO remedies range from zero damages to those that are \textit{strictly less} than “substantially equivalent to the level of nullification or impairment” – which WTO jurisdiction has interpreted to be roughly equivalent to the reliance damages. This is so, because damages are payable \textit{only} if the victim sues and the injurer subsequently is found guilty of the violation. The remedy that the injurer expects to incur is strictly smaller than the reliance damages payable under the formal opt-out mechanisms. This is so because he factors in (a) the probability that the victim may not go to court

\begin{footnotesize}
\begin{itemize}
  \item[40] The non-discriminatory feature of formal safeguard measures has been cited as one reason why it is infrequently used: Governments may prefer a more targeted instrument which can be directed at the country or set of countries that are the source of the contractual regret (e.g. Barton et al. 2006).
  \item[42] Throughout the GATT and the DSU, the \textit{same} remedy of withdrawal of \textit{substantially equivalent concession} appears as a countermeasure to legal and illegal nullification or impairment of previously agreed market access concessions (confer Art. 22.4 DSU, GATT Arts. XVIII.7.b, XIX.3.a, or XXVIII.3.a/4.d).
  \item[43] We can, however, think of a \textit{practical} reason for why the WTO drafters opted for lax punishment for violation of the WTO Agreement: The WTO framers may have been cognizant of their insufficient handling of the contractual incompleteness. Realizing that the current governance structure was insufficiently selective between good-faith and bad-faith clashes, the WTO’s founding fathers may have been afraid of enraging well-intentioned injurers by punishing them over and above the trade damage they caused. Instead, they opted for the stopgap solution of commensurate punishment for \textit{extra-}contractual behavior. To substantiate this conjecture, however, more historical research would have to be conducted.
\end{itemize}
\end{footnotesize}
at all, (b) the possibility that the injurer may actually win the litigation, (c) the case be suspended over legal formalities, (d) the certainty that trade damages are awarded prospectively, and (e) to the probability that the victim may refrain from enacting her retaliation award.44

By punishing legal and illegal behavior in the same fashion, the WTO contract to all intents and purposes establishes violation-cum-retaliation as the de facto default rule of non-performance: The lax punishment of Art. 22.4 DSU turns violation of the Agreement and losing a trade dispute into the de facto default flexibility instrument for any injurer. This fallback strategy at the same time is one of the most attractive escape tools (Mavroidis 2000). This is consequential, because it sets the benchmark for all other (formal and informal) escape remedies the WTO knows today. Moreover, it determines the power relationship in all settlement negotiations between injurers and victims: When bartering over voluntary compensation, no injurer will be willing to settle above his reservation utility, i.e. the expected cost of enduring retaliation.

To conclude this section on contractual escape and the reality of trade policy flexibility in the WTO: The efficient breach contract as the ideal contract in an uncertain world mandates equivalence between rules of ex post flexibility and entitlement protection: A functional system of non-performance must prohibit opportunistic ex post market-closing behavior, yet reap the opportunities looming by unforeseen regret contingencies. It also must compensate victims of post-contractual discretion so that their ex ante commitments do not get frustrated by ex post escape. Finally, the system must be able to protect signatories from illegal extra-contractual behavior.

The current WTO system of trade policy flexibility and entitlement protection misses that benchmark by a considerable margin: This is due to the fact that intra- and extra- contractual behavior are sanctioned in the same way. The floodgates to unchecked use of informal trade policy flexibility instruments in the WTO are thus gaping wide open. In consequence, the contemporary trade flexibility regime practically obliterates any de iure rule. While the de iure system of trade protection privileges victims and exacerbates ex post discretion of potential injurers, the de facto system seriously disadvantages victims of ex post discretion: It is not too difficult to see that in any given regret situation a rational injurer will always go for the very escape instrument which promises the “most mileage”, i.e. the least enactment costs, the lowest compensation payable to the victim, and the largest scope of application. With the informal flexibility tools VERs/OMAs, subsidies, domestic trade-related policies, AD or CvD actions at his discretion, any injurer can renege on the trade entitlement practically for free. Alternatively, the injurer can simply infringe upon the treaty at the price of losing a trade dispute and having to pay the WTO arbitrator’s estimate of damages, which – in expected terms – is strictly less than the amount owed to the victim under a formal safeguard. Violation-cum-retaliation is the ultimate de facto default rule of flexibility; no matter what the various WTO Agreements may say, the current flexibility regime is tantamount to a pure liability rule with the court’s interpretation of remedies equivalent to the harm done. This sets strong incentives for injurers to simply disregard the rules of the game.

44 An exception is the flexibility instrument of VER: Due to a mutual side-agreement, the victim does not file a complaint with the DSBM; instead, she bilaterally negotiates the size of the remedy. However, given that the injuring Member’s best alternative to a negotiated solution (his “reservation utility”) is the outcome of a trade litigation, he will not settle for a higher remedy than what he would expect to pay if he lost a trade dispute.
C. Literature review: Models of breach and remedy in the WTO and elsewhere

So far we stressed the intricate dynamism between cooperation in trade and trade policy flexibility. WTO scholarship is somewhat opaque as to how various trade flexibility mechanisms influence countries’ decisions to liberalize (ex ante) and to default (ex post). Starting with the works of economists like Dixit (1987), Bagwell and Staiger (1990), Riezman (1991), Hungerford (1991), Copeland (1990), or Kovenock and Thursby (1992), and international relations (IR) scholars like Downs and Rocke (1995), the GATT/WTO literature gradually moved away from an image of trade agreements as complete contracts in a stationary world. Acknowledging the presence of uncertainty concerning future events and agent behavior, trade agreements were modeled as incomplete contracts. This prompted trade scholars to examine the important role of trade policy flexibility mechanisms in trade agreements.

One strand of the literature (especially the contributions by Bagwell and Staiger 1990, Riezman 1991, Downs and Rocke 1995, Rosendorff and Milner 2001, Herzing 2005, Rosendorff 2005) focuses on the stability-enhancing role of contractual escape clauses. Modeling the WTO as a non-cooperative repeated prisoners’ dilemma, authors conjecture and/or prove that escape-clause induced negotiation equilibria Pareto-dominate those self-enforcement arrangements without contractual safety valves. Flexibility mechanisms were thus proven to successfully decrease the breakdown risk of the trade game and to add to the stability of the world trading system. Corollary findings are that the inclusion of trade policy flexibility instruments increases the circle of signatories and the probability of the conclusion of the trade accord in the first place.

While this literature is important in establishing a rationale for trade policy flexibility mechanisms in trade agreements, it fails on three accounts for the purpose of this paper: First, above models assume the presence of a single (exogenously given) escape mechanism, namely an unconditional liability-type opt-out that allows the injurer to partially withdraw from his commitments at his discretion. Hence, these approaches fail to discuss the crucial issue of rational escape clause design. In the same vein, assuming a pre-existing liability rule obviates scholars from having to take into consideration utilities and action spaces of potential victims of such backtracking behavior. Second, this literature stays largely elusive (or conjectural) about the dynamic implication which trade policy flexibility has for ex ante or future-round commitments (with the notable exception of Herzing 2005): Countries are assumed to always liberalize to full extent, no matter what the flexibility regime. Third, the trade-setting game is modeled as a non-cooperative game. Using the set-up of an infinitely repeated prisoners’ dilemma, co-operation is sustained by threatening to sanction defection with a “grim trigger response”, i.e. by once and for all resorting to the non-cooperative Nash past. Leaving aside the

45 See also Goldstein et al. (2000), Goldstein and Martin (2000), and Koremenos et al. (2001) for non-technical reasoning along those lines.
46 Victims’ need for compensation is often neglected. In Rosendorff and Milner (2001) and Rosendorff (2005), for example, the offer of compensation payments only serves as a signal, with which the injurer indicates his continuing willingness to cooperate in future periods.
47 In game theory, a non-cooperative game is one in which any instances of cooperation are self-enforcing. A game in which contracts are enforceable through social conventions, third parties or collective punishment, is a cooperative game.
question whether the threat of the grim trigger is in fact a credible punishment strategy or not, our approach is to interpret all contractual flexibility instruments as on-the-equilibrium-path strategies that are sustained by “the shadow of the grim trigger.”

Another strand of incomplete-contract WTO literature takes a slightly different route: Authors like Hungerford (1991), Kovenock and Thursby (1992), Maggi (1999), or Ludema (2001) model trade agreements as non-cooperative trade-setting games in the presence of an explicit enforcement authority, namely a trade court. Instances of ex post escape by an injuring Member are however neither fully observable, nor fully verifiable. Helpful as these models may be for giving economic explanations for trade courts (cf. Keck and Schropp 2007), they say little about the underlying dynamics of flexibility design. Above papers exogenously assume the pre-existence of trade courts which adjudicate conforming to the spirit of the contract. Dispute panels or the Appellate Body are believed to separate good-faith disputes (read: flexibility-enhancing non-performance) from opportunistic bad-faith protectionism. How the court overcomes possible information asymmetries, and why it does so better than parties themselves, is usually not discussed. In the same vein, the issues, why signatories may want to authorize an outside agent to complete an incomplete contract, and what concrete instructions for gap-filling and damage calculation the court is given by signatories, are neglected.

Taking trade policy flexibility to be cooperative games on the grim trigger-supported equilibrium path, a third branch of WTO literature assesses how the ex ante choice of escape provisions shapes the willingness of governments to liberalize trade in uncertainty-ridden trade accords. Studies like Sykes (1991, especially Appendix A and B), Ethier (2001), Roitinger (2004), and Bagwell and Staiger (2002b, chapter 6) model trade agreements as fully non-contingent tariff liberalization contracts and discuss the incentive-compatibility constraints posed by intra-contractual remedies. The authors examine the interrelationship of ex ante trade liberalization commitments and ex post exercise of trade policy flexibility. In so doing, these contributions integrate the concerns of victims of temporary escape by raising attention to the fact that adequate compensation for the victim is indeed crucial. All these approaches, however, consider only one single exogenously given liability-type flexibility mechanism, similar to Art. XIX GATT or violation of the Agreement, respectively. In addition, Ethier,

---

48 Klimenko et al. (2002) convincingly argue that an affected party’s threat to revert to non-cooperation is not credible, since such a strategy would be costly to apply: By returning to protectionism, the punishing party foregoes its share in the overall gains from international trade. Instead of “shooting itself in the foot”, the affected party has an incentive to return to the negotiation table in order to access the benefits of future cooperation. With this knowledge, injuring Members may have an incentive to deviate from the terms of the original agreement: They can heed the reasonable hope that their behavior will not trigger a mutual trade war, since every party involved has an incentive to resume cooperation.

49 The underlying rationale hereby is that the threat of resorting to a grim-trigger punishment and therewith of retreating to a highly protectionist past is tacitly supporting the system’s equilibrium path. Since hardly any Member should have a serious interest in returning to a pre-contractual non-cooperative Nash equilibrium (think of the consequences of the 1930 US Smoot-Hawley Act), “all actual trade disputes, punishments, and defiances [sic!] of [dispute settlement] rulings have been just parts of the equilibrium path” (Ethier 2001, p. 4). Neglecting concerns over system breakdown entails significant modeling convenience: Issues of non-performance, remedies, and ex ante commitment are not confined by the corset of a non-cooperative trade-setting game. Instead, they can be detached from enforcement and be modeled within a cooperative game setup, whose outcomes are supported by the threat of reverting to a grim-trigger punishment. This allows for more structure of the model via additional decisions and actions.

50 As to the difference between observability and verifiability (Schwartz 1992, p. 279): “[I]nformation is observable when it is worthwhile for the parties to know it, but the cost of proving it to a third party exceed the gains; information is verifiable when it is both observable and worth proving to outsiders.”
Roitinger, and Sykes model symmetrical contingency-revelation. Intuitively, if a market-access contingency is symmetrically revealed to both parties, the game has a trivial outcome: Since no injuring party has the possibility to misrepresent the true state of nature, the shock-affected party should be allowed to defer full contractual performance for exactly one period, given it pays commensurate damages to the victims. Aforementioned authors have thus shown that under symmetrical revelation of information a simple tit-for-tat result replicates the Pareto-efficient complete contingent contract.51

Symmetrical revelation of previously unforeseen market-access contingencies, however, is a very strong assumption. Adding some realistic flavor, Copeland (1990), Bagwell and Staiger (2005), Martin and Vergote (2004) and Herzig (2005, chapter 3) suppose informational asymmetry concerning the revelation of previously unforeseen contingencies (i.e. over the pressure on the injuring government to restrict trade). As is well known, asymmetrical information revelation invites strategic gamesmanship by the informed party (“moral hazard”). Modeling trade agreements as infinitely repeated tariff-setting games of the prisoners’ dilemma type, contributions in this category are also concerned with optimal levels of ex ante liberalization in the presence of trade policy flexibility options. However, again, these models presuppose one fixed liability-type escape design (violation-cum-retaliation, and AD-cum-retaliatory AD, respectively), and focus on revelation and payment schedules (sharing rules) that force the better informed party to tell the truth and to reveal its private knowledge in the face of the deterrent threat of a grim trigger strategy.

In WTO scholarship, our own approach most closely resembles the models of Feenstra (1987) and Feenstra and Lewis (1991). We, too, develop a cooperative one-stage game, where unforeseen contingencies are revealed asymmetrically. The two contributions are concerned with finding those combinations of ex ante liberalization level and transfer of rents that result in truthful contingency revelation by the injurer. Yet we aim at finding the proper flexibility design which induces optimal upfront commitments and efficient escape behavior despite the continuing presence of private knowledge of the size of protectionist shocks. In brief, while Feenstra and Feenstra/Lewis have only two choice variables (the optimal level of trade liberalization; compensation transfers), our model additionally endogenizes the ex ante choice of escape mechanism design.

Leaving the trade realm, the issue of breach and remedies in incomplete contracts has been discussed extensively in neighboring disciplines of international economics. In the field of industrial organization (IO), the link between ex ante commitment (in the form of “reliance investments”) and ex post non-performance has been discussed in detail. The basic insight is that the threat of escaping obligations can have a crucial impact on the ex ante commitment/investment behavior of the party which expects to become victim of that backtracking measure. This is the so-called hold-up problem.52

51 Note that symmetrical revelation of contingencies yields Axelrod’s classic result that tit-for-tat is the most viable retaliation strategy in an infinitely repeated prisoners’ dilemma (Axelrod 1984, Axelrod and Keohane 1986, Oye 1986).

52 Hold-ups occur in situations where one or both of two trading partners make relationship-specific reliance investments at the time of the contract conclusion. Reliance investments enhance the efficiency of this specific trade, but have considerable less outside value – they are partly sunk. These investments cannot be contractually fixed, either because they are non-verifiable by a court, or because future contingencies are imperfectly foreseen. Hold-ups then take place when one party (usually the one having made no prior investments) imposes contract renegotiations on the investing party. Threatening to cancel the deal, the injuring party can partially expropriate the quasi-rents generated from efficiency gains of the relationship-specific investment. Anticipating such opportunistic behavior, potential hold-up victims opt for underinvestment. This leads to an inefficient resource
While IO economists are concerned with the *ex ante* effects of contractual violations (or the threat thereof) in situations where enforcement is impossible, contract theorists from the law & economics (L&E) field have tackled the issue of the optimal design of remedial systems in incomplete contracts (the prominent “liability vs. property rule” debate; see Kaplow and Shavell 1996b at note 1, Dunoff and Trachtman 1999, and Krauss 1999 for literature reviews). Some L&E scholars discuss the differential impact of various non-performance provisions on *ex ante* commitment in contracts which can be enforced by an impartial court (e.g. Barton 1972, Shavell 1980, Rogerson 1984, Shavell 1984, Posner 1988, Edlin and Reichelstein 1996). Another strand of L&E compares the ease of use and the administration of liability and property rules in face of various post-contractual transaction costs (Calabresi and Melamed 1972, Ayres and Talley 1995a; 1995b, Kaplow and Shavell 1995; 1996a; 1996b).\(^{53}\)

As much as we would like to resort to the results generated by the IO and L&E literatures, and to simply apply them to trade agreements, such endeavor is not feasible: The hold-up literature in IO is unsuited as building block for a treatment of the WTO contract: Given that it assumes away enforcement of *extra*-contractual behavior, such assumption forecloses any sort of upfront agreement on flexibility mechanisms. The only thing which investing parties can do in a hold-up situation is to scale down their cooperative zeal according to the risk of being expropriated by the injurer.

The L&E literature on enforceable contracts is more suited to our needs, but cannot be copied-and-pasted, either: First, with the exception of Edlin and Reichelstein (1996) game setups featuring *bilateral* commitment/investment are not described in the literature. Second, contingencies are routinely assumed to be revealed symmetrically (the notable exception being Shavell 1980). Third, most contributions neglect *ex post* transaction costs and assume costless renegotiations. Qua Coase Theorem this assumption renders any discussion of trade policy flexibility design futile (cf. footnote 21 above). Fourth, virtually all models we are aware of suppose *pre-assigned roles*: Every player knows *ex ante* whether s/he is going to be victim or injurer (s/he just does not know the size of her/his regret and damage from non-performance, respectively).

The model which we are going to introduce in a minute most closely resembles Shavell (1980). However, we extend and modify the author’s approach in three important respects: Whereas Shavell assumes away any possibility of renegotiations, we shall integrate a property rule of flexibility in our considerations. Second, while Shavell models pre-assigned roles, we presume a “veil of ignorance”, behind which no player knows his future role with accuracy. Finally, whereas Shavell examines instances featuring unilateral investment, we assume that two players engage in bilateral *ex ante* commitment.

---

\(^{53}\) The liability-vs.-property dispute plays an important role in two other L&E contexts which are less relevant here: First, in the debate of *pre-transactional* allocation of property rights, where scholars are concerned to whom property rights shall be assigned in the face of transaction costs, and how to design efficient rules on the “taking of property” (injunction). In economics this issue features under the rubrics “residual rights assignment”, “comparative institutions” and “boundary of the firm” (see Grossman and Hart 1986, Hart and Holmström 1987, Hart and Moore 1988). Second, liability-vs.-property rule is debated in the literature on *prescriptive jurisdiction*: L&E scholars discuss the design of court-ordered rules of default in instances where contracting parties have failed to specify their secondary rules of entitlement *ex ante* (e.g. Ayres and Gertner 1989, Johnston 1990, Ayres and Gertner 1992, Craswell 1999).
D. A model of trade policy flexibility design in trade agreements

We assume two risk-neutral, self-interested, and reasonably rational players representing two symmetrical countries who negotiate a trade agreement behind a veil of ignorance. They design a governance structure for a single-transaction, one-shot tariff-setting game (or, equivalently but more realistically, a series of unrelated future periods). Enforcement against defection is safeguarded exogenously. Uncertainty is introduced in the form of a completely unexpected (exogenous) political support shock ($\theta$) of a finite size and known distribution. The trade agreement, hence, is organized as a completely non-contingent contract that consists only of a description of the level of market access concessions (we assume tariff cuts for simplicity), and the agreed-upon remedial regime (consisting of both an escape- and a remedy rule).

D.1 The economy

Consider a world with two symmetrical countries that are identical, except for their endowments in specific factors. Each country consumes three goods $x$, $y$, and $z$, whereby $z$ is the numéraire good with unit price of 1. On the supply side the home country ($H$) is relatively well endowed with the specific factor used for the production of good $x$, while foreign country $F$ (marked with a *) has a comparative advantage in producing good $y$. The Heckscher-Ohlin theorem implies that $H$ will export $x$ while $F$ will export $y$.

In accordance with the now standard insights in the field of endogenous tariff policy (cf. Rodrik 1995, Grossman and Helpman 2001 for comprehensive surveys) we assume self-interested policymakers who are in charge of formulating domestic trade policy. To that end, each policymaker (or government) has a single instrument at her disposal – a tariff on imported goods ($t$ and $t^*$, respectively). Government utility is assumed to be additively separable in the arguments of consumer surplus (CS), import-competing producer surplus ($\pi_m$), export producer surplus ($\pi_x$), and tariff revenues ($T$).

Policymakers’ utility shall be defined as

\begin{align*}
\text{Home: } W(t, t^*) &= CS(t, t^*) + \alpha IT_m(t) + \alpha^* IT_m(t^*) + T(t) ; \\
\text{Foreign: } W(t^*, t) &= CS^*(t, t^*) + \alpha^2 IT_m(t^*) + \alpha^* IT_m(t) + T^*(t^*) .
\end{align*}

whereby $\alpha, \alpha^* > 0$ are given exogenously and denote the political-economy weights that governments attach to firms’ profits.\(^54\)

---

\(^54\) These “politically realistic objective functions” (Baldwin 1987) capture the idea that government officials are politically motivated. They have specific domestic redistribution objectives. Governments are assumed to maximize some weighted average of general welfare (a measure that presumably safeguards votes) and special interest group-welfare (which upholds political support by special interests). The relative weight that general welfare and producer welfare assume in each policymaker’s calculation eventually depends on the personality of the respective policymaker and on the specific context at hand (Grossman and Helpman 1994).
D.2 Reason for contracting

Each country’s payoff, hence, is a function of its own tariff $t$ and the foreign tariff $t^*$, i.e. $W = W(t, t^*)$. There exists a best-response function (BR).

$$t_{ar}(t^*) = \arg \max_t W(t, t^*)$$

which allows each country to choose its tariff as a best response to the other country’s tariff level.

A common way of modeling international trade policy is to regard the tariff-setting interaction between two countries as a prisoners’ dilemma (Johnson 1953, Dixit 1987, Bagwell and Staiger 1990, Grossman and Helpman 1995, Bagwell and Staiger 1999; 2002b, Ethier 2004a, Ethier 2004b; 2006): In the beginning of every period and before trade takes place, each country can choose between applying the previously agreed cooperative import tariff and deviating with an optimal defection tariff. The following payoff structure results:

- $W(t_C^*, t_C^*) = W_C$ is each player’s per-period payoff in the Cooperative case;
- $W_D(t^*_{BR}(t_C^*), t_C^*) = W_D(t^*, t_C^*)$ is the payoff from unilateral Defection;\(^{55}\)
- $W_S(t_C^*, t^*_{BR^*}(t_C^*)) = W_S(t_C^*, t_D^*) = W_S$ is the “Sucker’s payoff” if the other party defaults on a cooperating country;
- $W_N(t^*_{BR^*}(t_{BR}^*)), t^*_{BR^*}(t_{BR}^*)) = W_N(t_N^*, t_N^*) = W_N$ is the non-cooperative Nash payoff that exists in the absence of any cooperative trading agreement.

The normality assumptions underlying the above payoff matrix can be summarized as:\(^{56}\)

1. $W$ is monotonously falling in $t^*$. This property is intuitive and warrants no further elaboration.
2. $W$ increases in $t$ for $t < t^*_{BR}(t^*)$, while it decreases in $t$ for $t > t^*_{BR}(t^*)$. Hence, for any given $t^*$, there is a unique $t$ maximizing $W$:

$$\frac{\partial W}{\partial t} \bigg|_{t^*_{BR}(t^*)} > 0 \quad \text{and} \quad \frac{\partial W}{\partial t} \bigg|_{t^*_{BR^*}(t^*)} < 0 \quad \text{(3)}$$

3. In case of mutual cooperation we assume that there exists a unique tariff level $t^*_C < t_N^*$ that maximizes $W_C(t^*, t^*)$. $t^*_C$ is defined as the cooperative tariff level which maximizes policymakers’ political welfare in a stationary world. Any trade liberalization beyond $t^*_C$ is welfare-decreasing even under cooperation. $t^*_C$ is important, since it indirectly affects the gains to be had from defecting ($W_D$), as well as the sucker’s payoff from being defected against ($W_S$).

4. There exists a tariff level $t' < t_C^*$ such that $W_C = W_N$ for $t_C^* = t'$. The argument hereby is more of a political than of an economic nature: Conceding to overzealous trade liberalization may induce political unrest by those special interest groups or parts of society that stand to lose from such policy (workers, import-competing industries, investors). In such a case a government may be better off without a trade liberalization accord at all. Hence, in our setting $W_C > W_N$ if and only if $t' < t_C^* < t_N^*$:

\(^{55}\) We assume that setting a tariff different from $t^*_C$ is regarded as a deviation from previously agreed obligations. A country will therefore always engage into optimal defecion. Its choice then becomes in fact binary between applying $t^*_C$ and $t_D^*$; partial performance (or partial defection) is not a rational option.

\(^{56}\) Slopes, curvatures and maxima of $W_{C,D,N,S}$ are described analogously in Herzing (2005, pp. 71).
\[
\frac{\partial W^C}{\partial t} > 0 \quad \text{for} \quad t < t_C^{\text{opt}} \quad \text{and} \quad \frac{\partial W^C}{\partial t} < 0 \quad \text{for} \quad t > t_C^{\text{opt}}.
\]

5. Furthermore, we assume that welfare gains from cooperation are increasing at a decreasing rate if 
\[t^C < t_C^{\text{opt}}\], and decrease at an increasing rate for 
\[t^C > t_C^{\text{opt}}\].
\[
\frac{\partial^2 W^C}{\partial t^2} < 0 \quad \text{for} \quad t < t_C^{\text{opt}} \quad \text{and} \quad \frac{\partial^2 W^C}{\partial t^2} > 0 \quad \text{for} \quad t > t_C^{\text{opt}}.
\]

6. A crucial assumption is that \(W^S\) increases at a decreasing rate in \(t^C\) for \(t^C < t^N\). Being cheated becomes less costly for the victim for lower levels of cooperation and thus for high realizations of \(t^C\).

7. Finally, \(W^D\) decreases in \(t^C\): The lower the cooperative tariff, the more one-time defection (“hit-and-run”) pays off. The increase in \(W^D\) is equal to that in \(W^C\) at \(t^C = t^N\) (where \(W^D\) and \(W^C\) are tangent), and unambiguously stronger as \(t^C\) falls below \(t^N\). Letting \(t^C\) decrease below \(t^N\) thus leads to a strong monotonous increase in \(W^D\).

For convenience, we adopt Herzing’s measure for the reciprocal level of trade liberalization, signified by the variable \(\tau\).

Defining
\[
\tau = \frac{W^D - W^C}{W^D - W^N},
\]

it is easily shown that \(\tau \in ]0, 1[\). Whenever tariff cooperation is approaching the non-cooperative Nash-level (which does not necessarily equal autarky), tariff liberalization is minimal. In other words, 
\[\lim_{t^C \to t^N} \tau = 0\]. A low cooperative tariff \(t^C\) corresponds to a high value of \(\tau\), and thus to a high degree of trade liberalization. Given that \(W^C_{\text{opt}} > W^N\) always holds, the above expression is bound below 1: Whenever \(t^C\) is only marginally larger than the lower boundary \(t^*\), the cooperative level of welfare \(W^C\) is only marginally higher than \(W^N\) as well. Thus, for \(t^C \to t^*\) the expression for \(\tau\) approaches 1 in the limit.

Figure 1 summarizes our normality assumptions:

---

57 Note that in a stage-game with symmetrical players, the expected per-period payoffs are identical for both signatories. We can thus drop the (*) representing the foreign country \(F\), and reduce our examination to one single signatory; the other Member’s payoffs are equivalent.

58 When \(t^C = t^N\), the expression is not defined, yet for trade liberalization infinitely bigger than Nash tariffs, the numerator is close to zero (defection is not very profitable), while the denominator is strictly bigger than zero.
Figure 1 plots a country’s welfare level against cooperative tariffs $t^C$. It is easily shown that the relevant range of cooperation is given by those tariff levels yielding Nash-superior payoffs, i.e. $t' < t^C < t^N$. In this range, reciprocal tariff liberalization is between 0 and 1.

In the relevant cooperation range, the following relationship holds: $W_D > W^C > W^N > W^S$ – the classical set-up of a prisoners’ dilemma. A prisoners’ dilemma has one Nash equilibrium – mutual defection is the dominant strategy. A contract, defined as “enforceable mutual commitment over time” (Craswell 1999, p. 18, Dunoff and Trachtman 1999, p. 30), can remedy this inefficient situation. Under pain of rigorous enforcement (e.g. in the form of the grim-trigger strategy or a collective punishment regime), the two symmetrical countries promise to commit to mutually beneficial trade cooperation and to stick to the rules of the game. In return, both players will receive higher cooperative per-period benefits generated from the stage-game.

**D.3 The nature of uncertainty and stages of the game**

So much for trade cooperation in a stationary world without imperfections. Things get more complicated, once we add realism by considering a non-stationary environment. In our model we are particularly concerned with three kinds of imperfection that non-stationarity entails:

(i) **Veil of ignorance.** We concur with Ethier (2001, 2002), Sykes (1991) or Lawrence (2003) that trade agreements, such as the WTO contract, are concluded behind a “veil of ignorance” (Rawls 1971). Reasonably rational countries negotiate the WTO as a long-term commitment in a repeated-interaction setting. Thereby, they cannot foresee (or time is too valuable) the composition of their comparative advantage, their role as injurers or victims, the identity and significance of acceding countries, and generally how future states of the world will impact their well-being.\(^{59}\) We model the

---

\(^{59}\) Readers may have doubts as to how realistic the concept of a “veil of ignorance” really is in trade agreements. Yet consider the following: The GATT as one of the fundamental pillars of today’s WTO system was founded in April 1947. Among its original signatories were Burma, Ceylon, the Republic of Cuba, the Czechoslovak Republic, Lebanon, Southern Rhodesia, and Syria. When signing the Agreement none of the signatories could foresee each others’ economic role six decades down the road. Czechoslovakia does not exist any more, other countries took quite different economic trajectories due to political upheavals (Zimbabwe, Burma, Cuba, Syria,
veil of ignorance in the following manner: At the time of the conclusion of the contract the two
governments of H and F anticipate a 50 per cent chance of experiencing regret and thus assuming the
role of victim or injurer.

(ii) Unforeseen contingencies and incompleteness of the contract. We assume complete uncertainty
of the future: Signatories cannot foresee (cannot put into watertight contract language) future
environmental contingencies or anticipate any trade-related domestic policy that trading partners may
come up with. And even if they are in a position to grasp such future contingencies, it is impossible for
policymakers to predict the complex domestic reverberations that such shocks will bear on their
personal political support functions. Uncertainty is thus introduced in the form of a completely
unexpected exogenous “political support shock” (θ) of a finite size.60

Unforeseen contingencies, or shocks, have the following characteristics:

• A shock is always “unwelcome”, i.e. unleashes a desire for a protectionist reaction in the affected
party. In other words, for a given τ the partial derivative WC,τ < 0 for any realization of θ. Another
way of seeing this is to assume that the political shock always harms import-competing industries
who then exercise political pressure on the self-interested policymaker; pro-export shocks are either
absent or do not stir a liberalizing reaction by pro-export lobbies. Formally, in a shock-ridden, non-
stationary world government utilities take the form:

\[ W(t, t^*; \theta) = CS(t, t^*) + (1 + \theta) \alpha \Pi_x(t) + \alpha \Pi_m(t^*) + T(t) \]  

(7)

• θ is a random i.i.d. shock of a commonly known probability density function \( p(\theta) \).61 θ
is finite and \( \in [1, \Theta] \).

• Shocks are idiosyncratic to each player, and are temporary in nature: They only affect one
signatory for a single period. The player affected by the shock will experience regret, and
consequently takes the role of the injurer.62

• For reasons of tractability, we assume that the best-response function of the defecting party is
unaffected by the size of the exogenous shock. In other words, \( t^D \) is independent of θ:
Independently of the realization θ, the injuring country will always apply the optimal defective
tariff (cf. footnote 55 above). This is a somewhat simplifying but convenient property that also

• The exogenous political support shock θ is assumed to affect the injuring country’s payoff
functions directly, but to leave the best-reaction functions of its trade partner unaffected. This is to
say that θ influences the injurer’s size of payoff, as well as its choice to default or not, but it leaves
the victim’s payoffs \( t^{D, I, C} \) essentially unchanged. The victim’s welfare is however affected by the
injurer’s choice of action (see Herzing 2005, p. 74 for an in-depth explanation).

---

Lebanon). According to economic dogma, completely rational policymakers would have had to take into
consideration developments of this sort, and discount them accordingly.

60 On the nature of political support shocks, see e.g. Rosendorff and Milner 2001, p. 832, Rosendorff 2005, p.

61 i.i.d. stands for “independent and identically distributed random variables”. In probability theory, a sequence
of random variables is i.i.d. if each has the same probability distribution as the others and all are mutually
independent. \( p(\theta) \) is exogenous, differentiable, and stable over time.

62 This property allows us to model global shocks that affect both countries simultaneously as separate incidents
– without loss of generalization.
(iii) Post-contractual transaction costs: Opportunity costs of renegotiation. When deciding on their preferred trade policy flexibility design, reasonably rational signatories must take into consideration the transaction costs that unfold after the conclusion of the contract. Importantly, tariff renegotiations involve significant opportunity costs: While countries are in the process of renegotiating, the surplus from seizing regret cannot be reaped. For simplicity (but not unrealistically), we will assume linear (iceberg) opportunity costs of time; every period that lapses without agreement reduces the gains from non-performance by a fixed amount. We assume that in the face of a protectionist shock in the injuring country, a speedy policy response will generate a bigger relief to the affected industry or societal group. The longer the renegotiations drag on, the larger the (political) damage in the injuring country. This is tantamount to saying that a slow renegotiation response implies less distributable efficiency gains from non-performance.\(^63\),\(^64\)

Figure 2 represents the stages of the game and locates the imperfections that non-stationarity entails:

![Figure 2: Stages and constraints in the trade game](image)

In stage 1 \((t_1)\), two self-interested policymakers as lawful representatives of their respective countries agree to succumb prisoners’ dilemmatic beggar-thy-neighbor policies, and to reap efficiencies through mutual trade cooperation. To that end, the signatories conclude a contract behind the veil of ignorance. Anticipating uncertainty over future contingencies and trade policy behavior, signatories enact an

\(^63\) Schematically, iceberg costs of renegotiation may be illustrated in the following way

![Image of iceberg costs](image)

The surplus from reacting to a shock \((\Phi)\) shrinks linearly over time as soon as renegotiations begin at a period \(t_1\), until at one point \((T)\) the non-performance surplus is zero.

\(^64\) Note that we will be neglecting costs of arbitration (quantification costs in case of a liability rule of escape), and litigation costs pursuant to both a LR and a PR. Technically, arbitration and litigation costs can be integrated as fixed costs. This, however, would not dramatically change our results.
“optimally indefinite contract”. A trade accord of this kind is a fully non-contingent contract that only consists of three unambiguous rules: The first provision is a substantive market access clause mandating parties to liberalize to the extent mutually agreed upon. The second clause is a specific flexibility instrument that specifies how parties may default from the contractually specified performance. The third rule prescribes how defection from these two rules shall be sanctioned.

Depending on the quality of the accorded trade policy flexibility mechanism, signatories in stage 2 negotiate the level of mutual commitments, i.e. depth of market access concessions. The pertinent choice variable is the reciprocal tariff level \( t (\tau^*) \). Trade liberalization commitments have to be reciprocal, which manifests in a binding level of \( \tau \).

After the level of cooperation is fixed, the performance phase of the contract begins with stage 3. In repeated rounds of tariff-setting the countries start interacting according to the terms of the contract. At some point in time \( (t_3) \), incertitude in the environment is resolved, and a single one-period shock occurs which is private information to every affected party. Being exposed to a shock may cause regret in the country that turns out to assume the role of the injurer.

In stage 4, depending on the terms of the initial agreement, an affected government decides whether to react to its political support shock by delivering as the contract stipulates, or by defaulting. If a rule of inalienability is in place, or if the shock was too insignificant to warrant expected compensation payments, the injurer will perform as promised. Whenever the contract provides for a LR and thus allows for ex post discretion, the injurer may default from previously agreed levels of performance at his own initiative. This triggers the contractually specified remedies in stage 5. Alternatively, the injuring Member may initiate a lengthy renegotiation phase aimed at buying off the victim’s right to demand performance as prescribed. As long as no agreement is reached, however, the contract must be performed as promised. Renegotiation takes the structure of a sequential request-offer bargaining

---

65 An optimally indefinite contract (Charny 1991, Cohen 1999) is the exact opposite of a comprehensive contract – but need not be any less efficient a substitute for a CCC. “In a world in which contract formation is costly and adjudication costless, a perfectly indefinite agreement, rather than comprehensive Arrow-Debreu bargains becomes the ideal contract”, (Masten 1999, p. 33).

66 Note well that we assume a functioning enforcement system to sanction extra-contractual behavior. The threat of resorting to a grim-trigger strategy and consequently of retreating to a protectionist past deters WTO Members from defecting from the previously agreed rules of the game (see our discussion at footnote 49 and accompanying text). Supposing rational behavior of the signatories, we assume that each stage-game is sustained by the threat of the grim trigger and that all events of non-performance are on-the-equilibrium-path actions, but not defection against the rules of the game. In other words, applying the optimal defection tariff \( t_D \) coupled with allocation of ex ante agreed compensation is not considered out-of-equilibrium (extra-contractual) behavior. It counts as legitimate excuse or escape from previously made tariff commitments and hence does not result in grim-trigger punishment.

67 For expositional convenience stages 1 and 2 are separated here. We think it is conceptually apposite to treat the negotiations concerning the contract’s institutional design and the depth of commitments as temporally separated instants: Parties must agree on fundamental contracting principles before deciding on their mutual reduction of trade barriers.

68 Since we assume symmetrical players, tariff commitments by the two players are equivalent. Were we to deal with asymmetrical players, the mutual tariff liberalization level \( \tau \) would be determined by the most reluctant liberalizer (cf. Ethier 2004a): A mutual deal between asymmetrical parties is struck at that level of cooperation where the first contracting party hits its participation constraint (Bagwell and Staiger 2002a, chapter 4). This constraint is then binding for all parties, since any deviation from that cooperation level necessarily brings Pareto-deterioration for at least one of the signatories. This is without loss of generalization, because the results of our model are not driven by the assumption of symmetry of players.
game with asymmetrical information. Renegotiations are resolved in a sequence of bargaining periods between \( t_3 \) and \( t_5 \): At \( t_5 \), the non-performance surplus is assumed to have “melted away”.

Stages 3 to 5 form the tariff-setting stage-game which is infinitely repeated. However, single tariff-setting stage-games remain unrelated.

\textbf{D.4 Notation summary}

For convenience we provide a summary of the notation we are going to use in the model. We define

\( \tau \) as the level of reciprocal tariff concessions on each country’s imported good, agreed to \textit{ex ante} by each player. \( \tau \) can be perceived abstractly as a “willingness to cooperate”.

\( \theta \) as an unforeseen (previously un-specified) contingency. \( \theta \) is a scalar that can be interpreted as an exogenous shock of some magnitude. The player affected by the shock experiences regret, and consequently takes the role of the injurer. The revelation of the shock is \textit{private knowledge} to the affected party;

\( p(\theta) \) as the probability density over \( \theta \);

\( B \) as \textit{breach set}, i.e. all those revelations of \( \theta \) that induce the injurer to escape from his contractual obligations and to decide for \textit{intra-contractual} non-performance; \{\theta | \text{the contract will not be performed}\};

\( W^C(\tau, \theta) \) as \textit{value} enjoyed by the \textit{injurer} if he gets hit by the exogenous shock, but \textit{performs} his contractual obligations as promised (Cooperation payoff);

\( W^D(\tau, \theta) \) as \textit{value} enjoyed by the \textit{injurer} if he gets hit by the exogenous shock and \textit{escapes} from his contractual obligations in the current stage-game (Defection payoff);

\( V^C(\tau) \) as \textit{value} enjoyed by the \textit{victim} if the injuring Member performs her obligations as promised (Cooperation payoff);

\( V^D(\tau) \) as \textit{value} enjoyed by the \textit{victim} if the injurer \textit{escapes} from his contractual obligations in the current stage-game (“Sucker’s payoff”);

\( V^N(\tau) \) as \textit{value} enjoyed by the \textit{victim} in the pre-contractual non-cooperative past (“Nash payoff”).\(^\text{69}\)

\( r \) as a contractually specified trade policy flexibility mechanism, consisting of a rule of breach (liability, property, inalienability) and a pre-sanctioned remedy damage \( d_r \), \( r \in [IR, z, rel, exp, PR] \), where \( IR \) stands for “inalienability rule”, \( z \) for “liability rule backed by zero damages”, \( rel \) for “liability rule backed by reliance damages”, \( exp \) for “liability rule backed by expectation damages”, and \( PR \) for “property rule”.

\( d_r \) as contractually agreed remedy (damage measure) payable to the victim, whereby \( r \in [z, rel, exp, PR] \).

\( \phi \) as breach surplus, i.e. the efficiency gains from non-performance: \( \phi = W^D - W^C \).

How \( \phi \) is shared between the players is a matter of the underlying design of the remedial regime, i.e. the trade policy flexibility mechanism chosen by the players.

\(^{69}\) We note that in the relevant cooperation range \( W^D(\tau, \theta) > W^C(\tau, \theta) \), and \( V^C(\tau) > V^N(\tau) > V^D(\tau) \).
D.5 The complete contingent contract as the benchmark in a non-stationary world

Inefficiency can occur for two reasons in our model: First, from inefficient levels of trade liberalization *ex ante*, and second from opportunistic breach decisions *ex post* (cf. footnote 26 above). As we have explained in footnote 20 and accompanying text, the complete contingent contract (CCC) is free from such imperfection. It is the contracting ideal in a non-stationary world – a hypothetical contract that signatories would write if no uncertainty whatsoever existed. The objective of such a perfectly contingent contract is to maximize the combined welfare of the two contracting parties along the Pareto-frontier (Shavell 1980). Under the CCC an injurer is only excused from performing *ex post* if, and only if, doing so strictly enhances the sum of payoffs to all signatories. In other words, the design of the CCC is equivalent to the plans of a benevolent social planner.

For a given level of trade liberalization, the combined objective function of the two signatory countries is:

\[
Z(\tau, B) = \left[ \int_{\theta}^{\Theta} V^{S}(\tau) p(\theta)d\theta + \int_{\theta}^{\Theta} V^{C}(\tau) p(\theta)d\theta \right] + \left[ \int_{\theta}^{\Theta} W^{D}(\tau; \theta) p(\theta)d\theta + \int_{\theta}^{\Theta} W^{C}(\tau; \theta) p(\theta)d\theta \right], \tag{8}
\]

whereby \( \theta \) stands for the lower level of shocks for which breach is observed. Equation (8) can equally be expressed in terms of breach and non-breach sets. Each time the privately observed shock \( \theta \) is above the threshold level \( \Theta \), the injurer is permitted excuse of performance. Hence,

\[
Z(\tau, B) = \left[ \int_{B} V^{S}(\tau) p(\theta)d\theta + \int_{\sim B} V^{C}(\tau) p(\theta)d\theta \right] + \left[ \int_{B} W^{D}(\tau; \theta) p(\theta)d\theta + \int_{\sim B} W^{C}(\tau; \theta) p(\theta)d\theta \right], \tag{9}
\]

where \( B \) is the breach set and \( \sim B \) the non-breach (i.e. performance) set. The terms in the first bracket represent the expected payoff of the *victim* country and the second bracket contains the expected payoff of the country subject to a shock. From the social planner’s point of view, given a certain shock, contractual performance is desirable as long as the combined welfare of cooperation yields bigger payoffs than defection. The higher the negative shock, the less attractive cooperation becomes, and the more likely is excuse from performance. We hence formulate

**Proposition 1:** A Pareto-efficient complete contingent contract exhausts all possible gains from trade. Whenever reacting to a large protectionist shock is apt to generate joint welfare gains, the CCC will prescribe breach, i.e. excuse from performance. Optimal breach *ex post* and optimal tariff liberalization commitment *ex ante* – and thus maximal welfare – are the result.\(^{70}\) More precisely, for any given level of trade liberalization \( \tau \), the Pareto efficient breach set equals \( B_{\text{opt}} \), where

\[
B_{\text{opt}} = \{ \theta \mid V^{S}(\tau) + W^{D}(\tau; \theta) \geq V^{C}(\tau) + W^{C}(\tau; \theta) \}. \tag{10}
\]

*Given the optimal breach decisions in the performance phase of a contract, a unique level of optimal upfront trade liberalization \( \tau_{\text{opt}} \) exists.*

**Proof:** The first part of *Proposition 1* is straight forward and does not require further explanations: Whenever pursuant to a shock (say, a balance-of-payment crisis) *ex post* escape yields efficiency gains, the CCC will prescribe excuse from contractual obligations. The optimal breach set is denoted by \( B_{\text{opt}} \). By definition, \( B_{\text{opt}} \) maximizes welfare for any given level of \( \tau \). The second part of *Proposition 1*

\(^{70}\) Note however that the maximand is not general (consumer) welfare, but the welfare of all contracting parties, viz. self-interested policymakers (cf. equations (1), (2) and (7)).
determines the optimal level of trade liberalization by differentiating the social planner’s common-welfare objective function \( Z(\tau, B_{\text{opt}}) \).

\[
Z(\tau, B) = \left[ \int_{\theta_{\text{opt}}}^{\theta_{\text{opt}}} V^C(\tau) p(\theta)d\theta + \int_{\theta_{\text{opt}}}^{\theta_{\text{opt}}} V^S(\tau) p(\theta)d\theta \right] + \left[ \int_{\theta_{\text{opt}}}^{\theta_{\text{opt}}} W^D(\tau, \theta) d\theta + \int_{\theta_{\text{opt}}}^{\theta_{\text{opt}}} W^C(\tau, \theta) p(\theta)d\theta \right].
\]

Knowing the threshold level \( \theta \) for which breach is preferred over performance, it is now possible to assign probabilities to these two events (breach, performance). Equation (11) can be rewritten in the following way:

\[
Z(\tau, B_{\text{opt}}) = \left[ (1 - \Pr(B_{\text{opt}})) V^C(\tau) + \Pr(B_{\text{opt}}) V^S(\tau) \right] + \left[ (1 - \Pr(B_{\text{opt}})) W^C(\tau, \theta) + \Pr(B_{\text{opt}}) W^D(\tau, \theta) \right].
\]

Constructing the first-order condition (FOC) yields

\[
\frac{\partial Z(\tau, B_{\text{opt}})}{\partial \tau} = \left[ (1 - \Pr(B_{\text{opt}})) V^C(\tau) + \Pr(B_{\text{opt}}) V^S(\tau) \right] + \Pr(B_{\text{opt}}) \left[ V^S_\tau(\tau) + W^D_\tau(\tau, \theta) \right] = 0.
\]

The result characterizes the optimal level of trade liberalization in a non-stationary world, \( \tau_{\text{opt}} \). Under the regularity assumptions of welfare, FOC (13) is sufficient to determine the global maximum: \( Z(\tau, B_{\text{opt}}) \) is continuous and differentiable. Since for each player there exists only one \( \ell^C_{\text{opt}} \) in the interval \([\ell', \ell^*]\), \( Z(\tau, B_{\text{opt}}) \) has a unique maximum in the trade liberalization level \( \tau_{\text{opt}} \).

\textit{Corollary 1: In a stationary world without shocks the welfare-optimal level of trade liberalization is higher than in a non-stationary world, where unforeseen shocks occur.}

\textit{Proof:} In a stationary world, rational players exhaust all benefits from trade, and fix the level of trade liberalization once and for all. No provisions for \textit{ex post} escape are necessary – deviation from the strict performance obligation is deterred by the grim-trigger threat. A benevolent dictator (or perfectly informed signatories) makes sure that

\[
\frac{\partial W^C(\tau)}{\partial \tau} = \frac{\partial V^C(\tau)}{\partial \tau} = 0.
\]

This condition is fulfilled when the reciprocal cooperative tariff is set equal to \( \ell^C_{\text{opt}} \), yielding a common tariff liberalization level of \( \tau^\text{SW}_{\text{opt}} \). Further trade liberalization beyond \( \tau^\text{SW}_{\text{opt}} \) would be detrimental to common welfare.

In order to compare the Pareto-optimal level of trade liberalization in a non-stationary world with the one in a stationary one, we solve equation (13) for \( V^C(\tau) \), the slope of the victim country’s cooperation payoff under the optimal level of \textit{ex ante} trade liberalization. Doing so yields

\[
\frac{\partial V^C(\tau)}{\partial \tau} = \frac{-W^C_\tau(\tau, \theta) + \Pr(B_{\text{opt}}) \left[ V^S_\tau(\tau) + W^D_\tau(\tau, \theta) \right]}{1 - \Pr(B_{\text{opt}})} > 0.
\]

71 Again, \( \tau_{\text{opt}} \) is the politically optimal tariff liberalization level from the point of view of self-interested policymakers. It is not the level of tariff concessions that maximizes general consumer welfare.

72 The superscript \( \text{SW} \) thereby stands for “stationary world”.

\[\text{– 28 –}\]
In line with our curvature assumptions, the first term in the numerator is unambiguously positive: Any shock is by definition unwelcome within the relevant cooperation range \( t' < t^C < t^N \). Hence, \( W^C_\tau(\cdot) = \partial W^C(\cdot) / \partial \tau < 0 \). The bracketed term in the numerator consists of two elements; the first being negative and the second positive. The sum \([V^S(\cdot) + W^D(\cdot)]\), however is strictly positive, because \( W^D(\cdot) > V^S(\cdot) \) holds: Under a CCC, breach occurs only if it is welfare-enhancing to do so, or formally \( V^S(\tau) + W^D(\tau, \theta) > V^C(\tau) + W^C(\tau, \theta) \). Yet this is only possible if the defection payoff increases faster than the sucker’s payoff falls, or in other words, i.e. if \( |V^S_\tau(\tau)| < |W^D_\tau(\tau; \theta)| \) holds.

\textit{q.e.d.}

**Interpretation:** In a \textit{non-stationary} world, an optimal policy requires the social planner to adapt his “investment” (i.e. the optimal level of \textit{ex ante} trade concessions), to the incidence of contractual breach. In the presence of shocks big enough to excuse performance, no gains from trade are reaped. The social planner thus must factor the probability of seeing less \textit{de facto} trade liberalization pursuant to a breach of one signatory. Consequently, he adapts the common \textit{ex ante} commitment by scaling down the cooperative zeal – the \textit{ex ante} cooperative “investment” simply yields less return than in a \textit{stationary} world. As a result, the optimal cooperative liberalization level is lower: \( \tau^{\text{opt}} \) lies somewhere to the left of the optimal liberalization in a non-stationary world (\( \tau^{\text{SW opt}} \)). Figure 3 illustrates:\(^{73}\)

![Figure 3: The optimal tariff liberalization level in a (non-)stationary world](image)

\(^{73}\) Note that in contrast to Figure 1 this chart plots tariff liberalization levels (and not tariff levels) on the x-axis. Hence, points further to the right represent \textit{higher} levels of liberalization (willingness to cooperate).
D.6 The individual’s maximization problem

The CCC may yield welfare-optimal outcomes, but signatories to a trade agreement are neither in the position to foresee (and write down) all future contingencies, nor do they usually feel the inclination to act like a benevolent social planner, especially when there exist imperfections in the market that they can capture on. Instead, it can safely be assumed that trade policymakers strive for maximization of their individual welfare and tend to disregard possible externalities that their behavior may ensue. Yet, what does the individual maximization problem of the self-interested policymaker look like - behind the veil of ignorance?

Proposition 2: Behind the Rawlsian veil of ignorance, the “efficient breach” property is the lynchpin of trade policy flexibility: Any agreed-upon trade policy flexibility mechanism that induces the injuring Member to engage in “efficient breach”, i.e. to default efficiently often, will automatically also result in the optimal trade liberalization level, and therefore replicate the outcome of the complete contingent contract.

Proof: Behind a veil of ignorance, every signatory party presupposes a 50% chance of assuming either role of injurer or victim. Hence, each party has an expected welfare of

\[
E[Y(\tau, B(\tau))] = \frac{1}{2} \left[ \int_{-B} V^C(\tau; \theta) p(\theta) d\theta + \int_{B} V^S(\tau; \theta) p(\theta) d\theta \right] + \\
\frac{1}{2} \left[ \int_{-B} W^C(\tau; \theta) p(\theta) d\theta + \int_{B} W^D(\tau; \theta) p(\theta) d\theta \right].
\]

(16)

Cancelling out \(\frac{1}{2}\), constructing the FOC, and solving for \(V^C(\tau)\) yields

\[
\frac{\partial V^C(\tau)}{\partial \tau} = -\frac{W^C(\tau; \theta) + \Pr(B(\tau))[V^S(\tau; \theta) + W^D(\tau; \theta)]}{[1 - \Pr(B(\tau))]} > 0.
\]

(17)

It is easy to show that the only difference between the FOC of the CCC (equation 15) and the individual policymaker’s maximization problem (equation 17) is the term \(B(\tau)\) (instead of \(B_{opt}\)), the breach set that is induced by a certain trade policy flexibility rule \(r\).

Any trade flexibility mechanism \(r\) apt to induce the injurer to breach efficiently often replicates the outcome of the contracting ideal of the complete contingent contract that a social planner would concoct. To see this, we simply replace \(B_{opt} = \{\theta \mid V^C(\tau; \theta) + W^C(\theta; \tau) \geq V^S(\tau; \theta) + W^D(\tau; \theta)\}\) for \(B(\tau)\), in equation (17) above. Doing so yields a result identical to equation (15), the FOC of the CCC.

q.e.d.

Proposition 2 is finding of some importance, since it demonstrates that signatories behind a veil of ignorance can effectively replicate the outcomes of the CCC, despite the presence of an incomplete contract. In order to design the “efficient breach contract”, signatories must devise a trade policy flexibility regime that induces efficient breach behavior in the injuring party. The remainder of this paper assesses different trade flexibility mechanisms as to their ability to generate this efficient breach property.
D.7 Inalienability rule as trade policy flexibility?

Some scholars perceive the WTO treaty as a *trade constitution* and as such as immutable. For proponents of this inalienability school of thought the WTO is a direct extension of national constitutions, aimed at securing the unconditional ab idance by basic economic property rights against protectionist backtracking or any kind of governmental opportunism.\(^74\)

According to inalienability proponents, WTO rules were deliberately conceptualized by Member States in a rigid manner, so as to tie the hands of current and future trade policymakers, who may otherwise be prone to fall prey to domestic protectionist pressures. Countries or rather peoples, engage in welfare-enhancing trade liberalization and demand freedom from any sort of governmental opportunism, which the inalienability perspective perceives to be a form of unlawful expropriation.\(^75\)

Constitutional rigidity upholds the predictability and stability of the system (as expressed in Art. 3.2 DSU) and advantages non-state actors who are normally under-represented in the domestic trade policymaking process (consumers, enterprises). Just as you cannot step back from constitutional obligations in unspecified circumstances, so the argument goes, it is equally pernicious to change the content of the WTO agreements in any way.\(^76\)

How does a rule of inalienability – a decision to prohibit trade policy flexibility all in all – jar with our model? On a positive note, an IR removes contractual uncertainty from consideration: Under a mandatory specific performance obligation the injurer’s breach set is the null set. On the negative side, however, we can already conclude that the resulting breach set is too small. Comparing inalienability with the optimal outcome under the CCC, it can easily be shown that there exist situations in which non-performance *ex post* would constitute welfare-increasing behavior.

A further interesting aspect is whether an IR increases or decreases the *ex ante* level of trade liberalization, relative to that prescribed by the CCC: A possible way of studying the behavior of contracting parties under increasing contractual stability is to take the limit of the individual FOC (15) and to approximate the probability of breach to zero.

\[
\lim_{\text{Pr}(B(\tau)) \to 0} \frac{-W^C_{\tau}(\tau; \theta) + \text{Pr}(B(\tau)) [V^S_{\tau}(\tau) + W^D_{\tau}(\tau; \theta)]}{1 - \text{Pr}(B(\tau))} = -W^C_{\tau}(\tau; \theta). \tag{18}
\]

---

\(^{74}\) Contributions to trade scholarship in the constitutionalist vein include Regan (2006); McGinnis and Movsesian (2000); Petersmann (1986, 2002, 2003), and Tumlir (1985). Hauser and Roitinger (2004, pp. 642) provide for an explanation and overview of the constitutional approach to the WTO.

\(^{75}\) Protectionist measures by domestic policymakers do not only give rise to economically inefficient production decisions, and therewith disadvantage domestic consumers and consuming industries. Moreover, they spur retaliatory actions (“sanctions”) by victim countries, which invariably are targeted at uninvolved domestic sectors that have not been at fault in any way. A retaliatory denial of market access then deprives innocent bystanders of their economic right to do business as previously consented to – a form of expropriation (see Charnovitz 2001 at note 159, Petersmann 2002; 2003, Bronckers and van den Broek 2005).

\(^{76}\) For another group of scholars, inalienability is just a codification of the most widely acknowledged peremptory international legal norm of *pacta sunt servanda*. Dunoff and Trachtman note (1999, p. 32): “Some international lawyers will reject the concept of efficient breach on a normative basis. They might argue that accepting the efficient breach hypothesis would threaten precisely the feature that renders treaties the ‘major instrument of international cooperation in international relations’ – the belief that treaties will be obeyed, even when contrary to the state’s immediate, short-term interest. Encouraging, through law, ‘efficient’ breaches of these treaties would undermine the fundamental rule of *pacta sunt servanda*, and likely render more difficult the possibility of sustained cooperation in an international community through treaty regimes”.
Given that $W^C(\cdot)$ by definition is negative (shocks are always unwelcome and unleash a protectionist desire in policymakers), the right hand side of equation (18) is unambiguously positive. As per Corollary 1 we know that under the CCC the slope of the expression $\frac{\partial V^C(\tau)}{\partial \tau}$ is also strictly positive.

Hence, in order to find out whether mutual trade concessions under the IR of flexibility (\(\tau_{IR}\)) are overzealous or underdeveloped, we have to check whether \(\tau_{IR}\) is to the right or to the left of \(\tau_{opt}\), the optimal level of tariff liberalization under a CCC (confer Figure 3). To do so, we compare the slopes of the FOC given by equation (15) and (18). In effect, we ask whether $-W^C(\tau;\theta) + \Pr(B_{opt})\left[V^S(\tau) + W^D(\tau;\theta)\right]$ is smaller or larger than $-W^C(\tau;\theta)$.

The decisive term hereby is $[V^S(\tau) + W^D(\tau;\theta)]$; the first term being negative, the second positive. In the proof of Corollary 1 we have already shown that the bracketed term must be positive under the CCC. Hence, we can conclude that trade liberalization $\tau_{IR}$ is strictly larger than $\tau_{opt}$, and that in Figure 3 $\tau_{IR}$ lies somewhere between $\tau_{opt}$ and $\tau_{SW_{opt}}$, the optimal trade liberalization level in a stationary world.

Under an IR and behind the veil of ignorance, signatories know that they will assume the role of the “victim” with a 50% chance every period. Cognizant of this fact, two opposing factors enter into consideration whenever a signatory chooses its ex ante tariff concessions: First, there is the risk of experiencing a protectionist shock and not being able to react adequately. Second, there is the possibility of becoming victim of a defective tariff and not be adequately compensated.

The reason why countries are tempted to engage in overzealous (and hence sub-optimal) ex ante levels of concession under an IR can now be explained as follows: Seen from the perspective of the injurer, trade liberalization is a risky endeavor. Given that an opt-out is not possible, countries hit by a protectionist shock cannot give in to their contractual regret. This should make them very wary of liberalizing too much. Future victims, on the other hand, calculate differently: Since the impeded injurer is barred from acting upon his regret contingency, a victim sees ex ante concessions as an investment with a contractually guaranteed rate of return. The victim thus behaves as if her ex ante investment (in the form of tariff concessions) were not risky at all – and is willing to liberalize up to the optimal level of concessions in a stationary world, $\tau_{SW_{opt}}$. She hence neglects the fact that even under the CCC ex ante commitment will not always produce efficiencies in instances of non-performance. Being well insured against all instances of breach on the part of the injurer, a victim’s private return to reliance exceeds the joint return. These concerns by the victim dominate the injurer’s considerations behind a veil of ignorance. As a consequence, ex ante tariff concessions are set excessively high.

In conclusion, we have shown that although a rule of inalienability generates contractual stability to future victims, the mechanism leads to inefficiently little ex post escape by injurers plagued by contractual regret. In addition, victims wrongly see tariff liberalization as a safe gamble and commit to too much trade liberalization compared to what would be optimal under a complete contingent contract. Anticipating such behavior, reasonably rational policymakers behind the veil of ignorance can therefore be expected to opt against an IR and for mechanisms that permit ex post escape.

\[77\] This kind of “overinvestment” is the contract-theoretical equivalent to the “tragedy of the commons” (cf. Hardin 1968).
D.8 Liability rules with varying damage measures as trade policy flexibility?

At the outset of contract negotiations, signatories may opt for a liability rule that allows the party experiencing regret (the injurer) to retreat from its obligations whenever he feels inclined to do so. Since the size of the exogenous shock is private knowledge to the concerned party, the only way to steer the actions of the injurer is by means of the choice of the remedy: The injurer’s decision to opt out of his obligations is clearly influenced directly by the size of compensation payments that he must indemnify the victim with. Rational behavior of the injurer requires that

\[
W^D(\tau; \theta) - d_r \geq W^C(\tau; \theta), \quad (19)
\]

which can be rewritten as

\[
W^D(\tau; \theta) - W^C(\tau; \theta) \geq d_r. \quad (20)
\]

The injurer’s participation constraint basically states that he will default if escape-cum-remedy payoffs are equal or bigger than performance payoffs, or in other words that the damage payment payable for non-performance must be strictly smaller than his efficiency gains from non-compliance. It is now possible to define the breach set in the following way:

\[
B(\tau) = \{ \theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq d_r \}. \quad (21)
\]

Behind the veil of ignorance signatories will choose the very level of trade liberalization which maximizes their expected level of welfare, taking into account the nature of the compensation scheme \(r\).

\[
E[Y(\tau, B(\tau))] = \frac{1}{2} \left[ \int_{B(\tau)} \left[ V^C(\tau, \theta) p(\theta) + \left[ \int_{B(\tau)} \left[ V^D(\tau, \theta) p(\theta) + d_r \right] d\theta \right] \right] d\theta \right] + \frac{1}{2} \left[ \int_{B(\tau)} \left[ W^C(\tau, \theta) p(\theta) + \left[ \int_{B(\tau)} \left[ W^D(\tau, \theta) p(\theta) - d_r \right] d\theta \right] \right] d\theta \right]. \quad (22)
\]

For every individual contractor, the damage payment \(d_r\) appears twice in the objective function – and eventually cancels out. However, since \(d_r\) is a main ingredient of the breach set \(B(\tau)\), it indirectly influences the level of trade liberalization that contracting parties deem optimal \(ex \ ante\). We will examine different LR remedies \(d_r\), and assess which one yields the closest outcome as defined under the perfect contingent contract.

D.8.1 Zero-damage rule

A first option to consider is the case of no damages measure \((d_r = 0)\). We integrate this option, because as our discussion in section B.3 showed, some de facto flexibility mechanisms in the WTO (mainly antidumping and anti-subsidy action as well as violation of the agreement) are essentially liability rules accompanied by zero-damages, if they are not challenged in court by the victim country, or won by the injurer.\(^78\)

\(^78\) See Schropp (2008, section 5.4). Trade scholars have called AD and CvD actions “ordinary protection with a good public relations program” (Finger and Zlate 2003), or “a poor man’s escape clause” (Hoekman and Leidy in Rosendorff and Milner 2001, p. 830), because they are easily enacted without having to compensate the victim(s).
Within the framework of our model any shock provokes regret in the country experiencing it: 
\[ \frac{\partial W^c(\tau; \theta)}{\partial \tau} < 0. \]
Without the deterrent of having to pay compensation, the injurer defects by optimally increasing his tariffs as soon as
\[ W^D(\tau; \theta) > W^c(\tau; \theta). \] (23)

The injurer will thus breach whenever he is subject to a shock. Since in every period one of the two countries is subject to a shock of finite size, the defective tariff \( t^D \) will be applied in every round of the tariff-setting game by whoever turns out to be the injurer.\(^79\) It is obvious that the breach set is too large and that escape occurs in each period. To see that, we insert the injurer’s participation constraint into the definition of the breach set (equation 21). The injurer’s decision to breach is driven by
\[ B(z) = \begin{cases} \theta \big| W^D(\tau; \theta) - W^c(\tau; \theta) \geq 0 \end{cases}, \] (24)
where \( z \) stands for the zero-damage instrument. It is now trivial to see that
\[ B(z) = \begin{cases} \theta \big| W^D(\tau; \theta) - W^c(\tau; \theta) \geq 0 \end{cases} \supset B_{opt} = \begin{cases} \theta \big| W^D(\tau; \theta) - W^c(\tau; \theta) \geq V^c(\tau) - V^S(\tau) \end{cases}. \]

A zero-damage remedy produces a breach set that is a superset of \( B_{opt} \). In other words, a remedy regime that provides no compensation payments induces the injurer to escape more often than is optimal. The country engages in over-breach and defaults inefficiently often, compared to the optimal breach set \( B_{opt} \). This is equivalent to saying that the injuring Member engages in non-performance at the occurrence of shocks of a lower magnitude than the CCC would permit default.

Given that the victim has no means of assessing the righteousness of the injurer’s claims to undergo “times of serious shocks”, she can only react by adjusting her ex ante commitments to the risk of being cheated. How then do signatories adapt their ex ante tariff concessions in anticipation of overzealous ex post breach by whoever turns out to be injurer? Consider the expected payoff behind the veil of ignorance:
\[ E\big[Y(\tau, B(z))\big] = \left[ \frac{1}{2} \int_{B(z)} [V^S(\tau; \theta) + d \int d\theta] \int [W^D(\tau; \theta) - d \int d\theta] \right]. \] (25)

Since under a zero damage rule escape happens with certainty in every period, equation (26) reduces to
\[ E\big[Y(\tau, B(z))\big] = \frac{1}{2} V^S(\tau) + \frac{1}{2} W^D(\tau; \theta). \] (26)

Constructing the FOC that each individual policymaker is faced with at the outset of contract negotiations yields
\[ \frac{\partial E[Y(\tau, B(z))]}{\partial \tau} = W^D(\tau; \theta) + W^C(\tau; \theta) = 0. \] (27)

---

\(^79\) The fact that AD and CvD actions are usually accompanied by some fixed costs (consultation, official investigation, and calculation of dumping margin/actionable subsidies) mitigates these outcomes a bit. In real life Members will not withdraw their concessions for any infinitesimal shock.
Since $W^D_\tau < 0$, we can rearrange equation (27):

$$V^S, (\tau) = W^D, (\tau; \theta)$$  \hspace{1cm} (28)

For equation (28) to yield a definite outcome, the marginal benefit from being defected against, $V^S_\tau (.)$, must equal the marginal benefit from escaping contractual obligations ($W^D_\tau$). This equality, however, only holds in the Nash solution, namely when $\tau^C = \tau^N$. If that were not the case, and if there existed a $\tau^C$ different from $\tau^N$ yielding bigger mutual payoffs, $\tau^N$ would be suboptimal. This, however, would defy the definition of the Nash equilibrium in the non-cooperative breach case.

In conclusion, a zero-damage rule leads to opt-out behavior by one signatory in every tariff-setting stage-game. Anticipating over-breach (relatively to the CCC), signatories behind the veil of ignorance are willing to commit to the non-cooperative Nash level, $\tau^N$. It goes without saying that the Nash level of liberalization is inefficient, simply because it is tantamount to the very trade liberalization concessions that exist in the absence of a contract.

**D.8.2 Restitution and reliance damages**

The compensation measure of *restitution* damages restores the *status quo ante* the *contract*. An injurer must reestablish the Nash level that persisted before the contract in the non-cooperative past. *Reliance* damages restore the *status quo ante* the *breach*. They compensate the victim for direct harm suffered, but leave aside indirect effects and foregone opportunities (such as transactional efficiency gains that would have accrued to in the case of normal contractual performance). We model a contract featuring a sequence of independent stage-games; therefore, both remedies are equivalent.\(^{80}\)

As we stated above (footnote 36 and accompanying text), WTO arbitrators arguably have interpreted the *equivalence* standard of Arts. XIX, XXVIII GATT and 22.4 DSU to be tantamount to reliance damages. But what is the effect of the reliance remedy on the behavior of WTO Members? In the model at hand, reliance (or restitution) damages can be defined as the difference between Nash- and the Sucker’s payoff:

$$d_{rel} = V^N - V^S (\tau) . \hspace{1cm} (29)$$

The participation constraint of the injuring country again is

$$W^D (\tau; \theta) - W^C (\tau; \theta) \geq d_{rel} ; \hspace{1cm} (30)$$

Reconfiguring equation (21), we define the injurer’s breach set:

$$B(\text{rel}) = \left\{ \theta \mid W^D (\tau; \theta) - W^C (\tau; \theta) \geq d_{rel} \right\} ;$$

$$B(\text{rel}) = \left\{ \theta \mid W^D (\tau; \theta) - W^C (\tau; \theta) \geq V^N - V^S (\tau) \right\} . \hspace{1cm} (31)$$

\(^{80}\) Reliance and restitution damages would differ if one country had made preparatory effort in the form of efficiency-enhancing reliance investments, or if we considered trade efficiencies to be accumulated over time.

\(^{81}\) Note that in our model the victim’s payoffs $V^N$, $V^S$, and $V^C$ are common knowledge to all players, since they are not directly dependent on the size of the (privately revealed) protectionist shock $\theta$. 

---

\(35\)
It is again easily shown that the reliance breach set is bigger than the optimal breach set under the CCC, since it holds that

- \( B(\text{rel}) = \{ \theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^N - V^S(\tau) \} \supset B_{\text{opt}} = \{ \theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^C(\tau) - V^S(\tau) \} \), and
- \( V^N > V^C(\tau) \) by the definition of a prisoners’ dilemma.

In other words, faced with reliance damages, the injurer engages in over-breaching – he defaults inefficiently often, compared to the optimal breach set under the CCC (\( B_{\text{opt}} \) is a subset of \( B(\text{rel}) \)). As said above, over-breaching is equivalent to opportunistic behavior by the injuring Member.

Let us now turn to the issue how signatories adapt their \textit{ex ante} tariff concessions to the expected over-breacht of future injurers. Replacing \( d_{\text{rel}} \) with \( VN - VS(\tau) \) and constructing the FOC of the individual policymaker, equation (22) rearranges into

\[
E[Y(\tau, B(\text{rel}))] = \frac{1}{2} \left[ \int_{d_{\text{rel}}} \left( V^C(\tau) p(\theta) d\theta + \int_{B(\text{rel})} [V^N(\tau) + d_{\text{rel}}] p(\theta) d\theta \right) \right] + \frac{1}{2} \left[ \int_{d_{\text{rel}}} \left( W^C(\tau; \theta) p(\theta) d\theta + \int_{B(\text{rel})} [W^D(\tau; \theta) - d_{\text{rel}}] p(\theta) d\theta \right) \right].
\]

Replacing \( d_{\text{rel}} \) by \( [VN - V^S(\tau)] \) yields:

\[
E[Y(\tau, B(\text{rel}))] = \frac{1}{2} \left[ \int_{d_{\text{rel}}} \left( V^C(\tau) p(\theta) d\theta + \int_{B(\text{rel})} [V^N(\tau) + VN - V^S(\tau)] p(\theta) d\theta \right) \right] + \frac{1}{2} \left[ \int_{d_{\text{rel}}} \left( W^C(\tau; \theta) p(\theta) d\theta + \int_{B(\text{rel})} [W^D(\tau; \theta) - VN + V^S(\tau)] p(\theta) d\theta \right) \right].
\]

Construing the FOC, we see that

\[
\frac{\partial E[Y(\tau, B(\text{rel}))]}{\partial \tau} = \left[ 1 - P(B(\text{rel})) \right] \left[ V^C, (\tau) + W^C, (\tau, \theta) \right] + P(B(\text{rel})) \left[ W^D, (\tau, \theta) + V^S, (\tau) \right] = 0.\]

\textit{Ex ante} trade liberalization concessions \( \tau_{\text{rel}} \) must be such that the above FOC (27) holds. As we have done previously, we now solve the equation (27) for \( V^C, (\tau) \):

\[
V^C, (\tau) = \frac{V^C, (\tau) + P(B(\text{rel})) [W^D, (\tau, \theta) + V^S, (\tau)]}{1 - P(B(\text{rel}))} > 0.
\]

How does the optimal \textit{ex ante} trade liberalization level under reliance damages (\( \tau_{\text{rel}} \)) compare with \( \tau_{\text{opt}} \), the globally optimal level of tariff concessions? Comparing equations (15) and (35) brings the answer: Since \( P(B(\text{rel})) > P(B_{\text{opt}}) \) holds, equation (35) is clearly bigger than (15). The slope of the victim country’s cooperation payoff under reliance damages is steeper than under the CCC, which is equivalent to saying that the \textit{ex ante} trade concessions under reliance are smaller than under the CCC contract.

Seen from the perspective of the victim, two aspects speak against overly liberal tariff concessions: First, victims are cognizant of overzealous breach by injurers. Every instant of breach means that no gains from trade are realized, and the expected value of the contract hence decreases. Second, a future
victim recognizes that the reliance damages which the injurer expends are smaller than the actual costs borne by the victim. Compensated by reliance damages, the victim would forgo efficiency gains from ongoing trade cooperation. Any efficiency loss from escaping the original contractual obligations will have to be shouldered by the victim alone, and she will not have a stake in any efficiency gain seized from non-performance, either. Thus, excessive escape behavior on the part of the injurer will leave the victim in a worse state than had he performed. Factoring into the calculation this opportunistic escape behavior by future injurers will bring down signatories ex ante willingness to cooperate behind the veil of ignorance.

In conclusion, reliance damages are not able to replicate the optimal outcomes of a CCC: First, ex post breach behavior on the part of injurers is suboptimal; second, ex ante commitments are lower than what would be efficient, because victims factor in both opportunistic breach and under-compensatory damages.

**D.8.3 Expectation damages**

Finally, we consider the possibility of expectation damages, a remedy solution suggested in the WTO realm by Mavroidis (2000), Sykes (2000), Schwartz and Sykes (2002), or Schropp (2005, 2007, 2008). Expectation remedies are equivalent to the replacement value that exactly makes the victim indifferent between the injurer’s performance and default – the remedy insures the victim against any dynamics that unfold ex post. The compensation payment $d_{exp}$ is defined by

$$d_{exp} = V^C(\tau) - V^S(\tau);$$

(36)

Under expectation damages, the participation constraint of the injuring country is given by

$$W^D(\tau; \theta) - W^C(\tau; \theta) \geq d_{exp},$$

or equally

$$W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^C(\tau) - V^S(\tau).$$

(37)

(38)

**Proposition 3:** Expectation damages induce efficient ex post breach decisions in injuring countries; the breach set $B(exp)$ is identical to the optimal breach set under the CCC, $B_{opt}$. This in turn safeguards optimal reciprocal trade liberalization concessions ex ante.

**Proof:** The first part of the proof is straightforward. Inserting participation constraint (38) into the breach set of the injurer, we see that

$$B(exp) = \{\theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq d_{exp}\};$$

$$B(exp) = \{\theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^C(\tau) - V^S(\tau)\};$$

$$B(exp) = \{\theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^C(\tau) - V^S(\tau)\} = B_{opt}.$$

(39)

Equation (39) shows that the expectation damage measure induces the same breach decision as the CCC, namely $B_{opt}$. 
The second part is basically a restatement of the proof we advanced for Proposition 2: Given the expectation remedy, the individual maximization problem of each policymaker is

$$E[Y(\tau, B(\text{exp}))] = \frac{1}{2} \left[ \int_{-\text{exp}}^{\text{exp}} V^C(\tau)p(\theta)d\theta + \int_{B(\text{exp})}^{\text{exp}} V^S(\tau)p(\theta)d\theta \right] + \frac{1}{2} \left[ \int_{-\text{exp}}^{\text{exp}} W^C(\tau; \theta)p(\theta)d\theta + \int_{B(\text{exp})}^{\text{exp}} W^D(\tau; \theta)p(\theta)d\theta \right].$$

(40)

Cancelling out $\frac{1}{2}$, constructing the FOC, and solving for $V^C(\cdot)$ yields

$$\frac{\partial V^C(\tau)}{\partial \tau} = -\frac{W^C(\tau; \theta) + \text{Pr}(B(\text{exp}))\left[V^S(\tau) + W^D(\tau; \theta)\right]}{1 - \text{Pr}(B(\text{exp}))} > 0.$$  

(41)

It is easy to show that the only difference between the FOC of the CCC (equation 15) and equation (41) of an individual maximization problem is the term $B(\text{exp})$. Acknowledging that $B(\text{exp}) = B_{\text{exp}}$, we see that equation (41) and (15) are indeed identical.

$q.e.d.$

The interpretation of this result is quite intuitive: Behind a veil of ignorance, signatories aim for a trade policy flexibility mechanism that mandates the injurer to internalize the externalities he produces, i.e. the costs of default. Rational cost-benefit calculations then induce injuring countries to withdraw from previously made concessions only in cases where \textit{ex post} escape is strictly welfare-enhancing. Receiving enough compensation to insure them against any \textit{ex post} protectionist backlash, future victims are willing to liberalize generously, yet they refrain from engaging in overzealous \textit{ex ante} liberalization (as they would do under an IR). Future injurers are equally prepared to liberalize optimally, since the expectation remedy allows them to seize future regret contingencies.

In conclusion, a liability rule accompanied by the remedy of expectation damages replicates the outcomes of the contracting ideal, the Pareto-efficient complete contingent contract: Only the expectation remedy possesses the incentive-compatible characteristics that effectuate efficient non-performance by the injurer \textit{ex post}, and optimal commitment by all signatories \textit{ex ante}.

\textbf{D.9 Renegotiations as efficient trade flexibility rule?}

Another possible way to cope with unforeseen contingencies is to introduce a property rule of flexibility, which mandates that regret be exclusively dealt with by means of \textit{ex post} renegotiations. Under a PR, a party wishing to temporarily modify the agreed-upon terms is free to negotiate a release from performance. No backtracking measure, however, can be enacted against the consent of the victim, or else the rules of the contract are violated and the enforcement mechanism kicks in.
A PR is the preferred escape option of the so-called compliance school of WTO scholarship. Compliance advocates concede that the WTO treaty is indeed an incomplete contract in a non-stationary environment (Sykes 2000, p. 347), but they fundamentally disagree with “rebalancing” proponents as to the form in which efficient ex post adjustment is to take place. Whereas under the precepts of a rebalancing system an injurer can adapt to changing circumstances unilaterally by enacting a general liability rule, compliance advocates beg to differ. They allege that a rule of renegotiation is not only a legal obligation, but also apt to yield fairer and more efficient results. Concretely this means that the only permissible manner of reacting to unforeseen developments, and to withdraw tariff concessions, is by way of renegotiations under the purview of Art. XXVIII GATT.

In order to appreciate how a PR of negotiations affects upfront cooperative behavior and ex post breach in our model, we have to slightly extend our game set-up. Modeling renegotiations as a simple sequential bargain, we model an alternating request-offer game of finite length, in which only the injurer knows with certainty the size of the renegotiation gains. Asymmetrical information over a common bargaining surplus is an important element here.

The two signatories bargain over the size of compensation payable by the injurer I in exchange for victim V’s waiver of demanding contractual performance. Importantly, the whole bargaining process is influenced by the fact that the distributable non-performance surplus (φ) is “melting away” over time. Injurer I enters into renegotiations with the clear intention of maximizing his share of the non-performance surplus. To that end, the informed injurer is tempted to engage in moral hazard, i.e. to understate the real size of the shock and to garble the actual non-performance efficiencies. The victim,
equally aiming to maximize her share of the spoils, will have to shape her compensation objectives according to what “type” of shock she believes the injurer was affected by.

We relegate the calculations in connection with the bargaining game to Appendix A and state directly our findings:

- A PR of flexibility produces inefficiencies for two reasons: First, negotiations break down with a positive probability and prevent welfare-enhancing escape from happening. Second, a property rule induces waste, because agreement is often reached with delay, so that non-performance efficiencies have partly “melted” away.

- Independently of which party makes the final offer, player $V$ can always expect to receive some spoils of the non-performance efficiencies. The victim country (in expected terms) can heed the hope to reap compensation payments over and above expectation damages.

- Consequently, the injurer (in expected terms) will have to pay the victim compensation payments higher than expectation damages. As we will show below, this has an effect on $I$’s breach behavior.

Using the results generated in Appendix A, we can see that, should the injurer choose to engage in renegotiations, his compensation payment for contractual relief in a three-period bargaining game is equal to

$$d_{PR} = d_{exp} + \frac{X}{2},$$

(42)

where $X > \theta$ constitutes the highest possible excess surplus of non-performance (cf. equation (A1) in Appendix A). Given this damage payment, we can proceed to construct the breach set of the injurer under a property rule:

$$B(PR) = \left\{ \theta \mid W^D(\tau, \theta) - W^C(\tau, \theta) \geq d_{PR} \right\}$$

$$B(PR) = \left\{ \theta \mid W^D(\tau, \theta) - W^C(\tau, \theta) \geq V^C(\tau) - V^S(\tau) + \frac{X}{2} \right\}.$$

(43)

Comparing equation (43) to the optimal breach set under a CCC (equation 11), we see that

$$B(PR) = \left\{ \theta \mid W^D(\tau, \theta) - W^C(\tau, \theta) \geq V^C(\tau) - V^S(\tau) + \frac{X}{2} \right\} \subset \left\{ \theta \mid W^D(\tau; \theta) - W^C(\tau; \theta) \geq V^C(\tau) - V^S(\tau) \right\} = B_{opt}.$$  

(44)

Thus, a PR rule produces a breach set that is a subset of $B_{opt}$. In other words, seen from the perspective of the injurer, engaging in ex post breach is inefficiently costly. The injurer “under-breaches”, i.e. he refrains from breaching in instances where the CCC would have mandated excuse. The level of shock necessary for profitable breach under a property rule of flexibility lies above the optimal shock-level defined by the CCC.

---

87 We characterize the outcomes of a three-period bargaining game as representative of an odd-period renegotiation, which is characterized by the victim making the final offer. Appendix A also describes the outcomes of an even-round bargaining game where the injurer makes the final bid. The results, however, are not fundamentally altered by changing from an odd to an even bargaining game.
At the time of the conclusion of the contract, that is behind the veil of ignorance, contracting parties have the following expected welfare function:

$$
E[Y(\tau, B(PR))] = \frac{1}{2} \int V^C(\tau)p(\theta)d\theta + \int [V^S(\tau) + d_{fa}p(\theta)d\theta] +
$$

$$
\frac{1}{2} \int W^C(\tau, \theta)p(\theta)d\theta + \int [W^D(\tau, \theta) - d_{fa}p(\theta)d\theta] .
$$

(45)

When inserting the results of a renegotiation game into equation (45), we must distinguish between instances of successful renegotiations (leading to efficient breach) and those where renegotiations break down inconclusively, and breach is refused by the victim:

$$
E[Y(\tau, B(PR))] = \frac{1}{2} \int V^C(\tau)p(\theta)d\theta + \int [V^S(\tau) + V^C(\tau) - V^S(\tau) + \frac{X}{2}] p(\theta)d\theta +
$$

$$
\frac{1}{2} \int W^C(\tau, \theta)p(\theta)d\theta + \int [W^D(\tau, \theta) - V^C(\tau) - V^S(\tau) + \frac{X}{2}] p(\theta)d\theta .
$$

(46)

The agreed level of \textit{ex ante} trade liberation maximizes the expected level of welfare. Constructing the first-order condition yields

$$\frac{\partial Y(\tau, B(PR))}{\partial \tau} = [1 - \Pr(B(PR))][V^C(\tau) + W^C(\tau, \theta)] + \Pr(B(PR))[W^D(\tau, \theta) + V^S(\tau)] = 0 .
$$

(47)

The optimal \textit{ex ante} trade liberalization level is chosen in a way such that the above FOC holds. How does this outcome perform in comparison with the optimal trade liberalization level $\tau_{opt}$? As before, we solve equation (47) for $V^C(\tau)$:

$$\frac{\partial V^C(\tau)}{\partial \tau} = \frac{W^C(\tau) + \Pr(B(PR))[W^D(\tau, \theta) + V^S(\tau)]}{1 - \Pr(B(PR))} > 0 .
$$

(48)

Just like under the optimal CCC contract, $V^C(\tau)$ is bigger than zero. Equations (48) and (15) however differ in the probability of witnessing breach. Since $Pr(B(PR)) < Pr(B_{opt})$, overzealous liberalization is the inevitable result. As is the case under an IR, future victims under a PR maximize their expected welfare and tend to ignore the negative externalities on global welfare (see footnote 77 and accompanying text for explanations).

In conclusion: A property rule of trade flexibility leads to under-breach by the injurer and to inefficiently high levels of \textit{ex ante} tariff liberalization.

\textbf{D.10 Summary of findings and caveats}

Our model of the WTO as a tariff-liberalization accord between self-interested policymakers has shown why “there does not exist a damage measure which is always Pareto superior to the expectation measure” (Shavell 1980, p. 483, emphasis in original). A trade policy flexibility mechanism designed as a liability rule backed by expectation remedies Pareto-dominates all other escape regimes (including a property rule). LR-cum-expectation damages is apt to replicate the outcomes of the unattainable complete contingent contract and therewith qualifies as the achievable first-best, the efficient breach contract.
Whenever signatories negotiate a trade agreement behind the veil of ignorance, prohibiting any *ex post* flexibility (mandating inalienability) and prescribing renegotiations leads to inefficiently little breach and overzealous *ex ante* liberalization. A liability rule accompanied by zero- or reliance damages gives rise to opportunistic over-breach by injurers. This not only brings down the mutual welfare level (hence constitutes Pareto-inferior behavior), but in addition discourages reciprocal tariff liberalization concessions that countries are willing to concede to *ex ante*. Only expectation damages unleash the optimal mutual commitment at the time of the conclusion of the contract, and provoke efficient escape behavior in the performance stage of the contract.

Figure 4 summarizes diverse effects on the *ex ante* trade liberalization decision that various trade policy flexibility regimes entail. The chart shows where the different trade policy flexibility mechanisms (IR, PR, zero-, reliance-, expectation damages) are located relative to the optimal levels in a stationary and non-stationary world ($\tau_{SW}^{opt}$ and $\tau_{opt}$, respectively).

It is easily seen that only the expectation remedy extracts the same *ex ante* tariff liberalization of signatories as the Pareto-efficient CCC would.

Some observations and qualifications seem apt when interpreting our model results:

1. Our model outcomes are very much driven by the assumption of the veil of ignorance: The two negotiating parties do not know with precision their future “role” as a victim or injurer following a protectionist shock. Examining similar games like ours, authors like Shavell (1980), Rogerson (1984) and Edlin/Reichelstein (1996) assume fixed roles. Doing so yields a significantly different set of results.\(^88\) Both approaches have their merits. Supposing fixed roles at the contracting stage is appropriate for modeling short-run strategies in trade agreements between heterogeneous countries. In reality, it is often clear which countries will be sensitive to political pressure as a result of trade liberalization.\(^89\) Our model is more suited to explain long-term, open-ended trade

---

\(^88\) Aforementioned authors find that expectation damages do not generate Pareto-efficient outcomes and instead may provoke overinvestment by the victim.

\(^89\) Think of China’s accession to the WTO: Chinese textile and shoe exports were anticipated to constitute a vital threat to some industrialized countries. These countries consequently pushed China into consenting to an extended phase-in period of free trade in those two sectors.
liberalization games between homogenous countries which trade many goods and services, and oftentimes have “knife-edge” comparative advantages.

2. The choice between a liability and property rule is not just one of the size of the remedies and the timing of breach: Installing a liability-type flexibility rule presupposes the existence of an impartial and competent arbitration instance that can assess trade damages incurred by the victim. Having at hand an arbitrator is connected to sunk-cost investments, which we did not integrate into our model. Also, we assumed away some real-life opportunity costs in connection with both an arbitration and renegotiation process, such as litigation costs, information gathering and processing costs, human capital costs, etc. (but see footnote 64).

3. In our model, assessing damages is straightforward, because the size of damages is common knowledge.\(^90\) In real life, however, calculating trade damages is a tough task (see e.g. Keck 2004, WTO 2005). Quantifying expectation damages is an even more daunting challenge. The reason is that no one knows with certainty how protectionist policies really affect victim countries.\(^91\) In order to examine the optimal trade flexibility regime under bilateral information asymmetry we would need to construct a trade model in which the injurer is ignorant of the implications of his escape, and the victim is ignorant about the true nature of the exogenous shock. We leave the formulation of a game of double-sided moral hazard to future research. Our conjecture, however, is that again a property rule is inferior in such a game: Under a liability rule, an independent WTO arbitrator would have to assess trade damages pursuant to an opt-out by the injurer. Doing so is neither cost- nor flawless.\(^92\) However, under bilateral information asymmetry, a flexibility rule of renegotiations may well lead to a “war of attrition”, in which both parties mistrust each other’s assertions, and consequently reject settlement offers.\(^93\) Wars of attrition take very long to resolve and hence are very costly.\(^94\)

4. We modeled the WTO as a tariff-reduction contract aimed at granting reciprocal market access to its Members. However, we acknowledge the fact that the WTO is not only about mutual market

---

90 The injurer does not adapt his opt-out behavior to the size of the shock and always defects optimally, thus causing a commonly known welfare loss of \(V^C - V^D\) to the victim (cf. footnote 55 above). Neglecting partial performance pursuant to the occurrence of a protectionist shock is a starkly simplifying assumption. Yet, we believe it to be without loss of generality: A more general payoff structure to include partial performance by the injurer would considerably increase the mathematical complexity of the model (in the form of an additional choice variable), but it would neither alter the underlying structure of the game nor the logic of the outcome.

91 Confer Schropp (2008at p. 324): “All trade agreements are inherently political deals designed and concluded by self-interested policymakers. Hence, the initial balance of concessions and the entire metric of the WTO is presumably profoundly political. Yet, arbitrators can neither observe nor measure the political harm done to a victim government by its trade partner’s unilateral policy adjustment. How can arbitrators ever claim to be able to calculate in tangible currency political expectation damages? How can they credibly assert to be able to re-establish a profoundly political balance of welfare if political expectation damages are unverifiable?”

92 Interestingly, Kaplow and Shavell (1996a, pp. 726) have shown that an independent arbitrator neither need be omniscient, nor operate flawlessly in order to be effective. As long as its judgment is neither completely incompetent, nor systemically biased, arbitrators can succeed in producing satisfying outcomes (on this account see also Rosendorff 2005).

93 Since both the regret contingency and the damage caused are private information to the renegotiating parties, both have a strong incentive to misrepresent the true state of nature: The victim will try to misconstrue her likely damage so as to appropriate as much as possible of what she thinks are the injurer’s gains from non-performance. The injurer has an incentive to misconstrue both the actual size of the exogenous shock and the extent of expected efficiency gains from non-performance.

94 Solution concepts for bilateral information are quite difficult (cf. Ordoover and Rubinstein 1986, Osborne 1985, or Henricks and Wilson 1985 for overviews of the literature). For our purpose at hand it suffices to say that wars of attrition incur severe efficiency losses (time costs) and opportunity costs of bargaining.
access and tariff reductions (see generally Pauwelyn 2006, Schropp 2008). Many more rights and obligations are being exchanged in the WTO, such as those concerning minimum standards (e.g. in the realms of intellectual property, investment, or international standardization), or procedural guidelines (e.g. timelines, transparency obligations, notification requirements). Not all our findings concerning trade policy flexibility carry over easily to each and every regulatory realm covered by the WTO. Whether flexibility concerning some notification requirement or a TRIPS-obligation shall be granted on the basis of a liability, property, or inalienability rule, this paper cannot answer.

5. Trade policy flexibility design in trade agreements is not only an issue of the form of escape (IR, PR, LR) and the intensity of remediation, but also about the form of remedy instruments: Whether compensation in the WTO should be payable to the victim in the form of monetary fees or tariff compensation, or whether it should be organized as trade sanctions (tariff retaliation, cross retaliation, collective sanctions) is not the topic of our paper. In our model, we have assumed monetary compensation for convenience only.

6. Finally, there is the issue of fairness: Is a trade agreement that features liability-cum-expectation damages a “fair” agreement, or does a property rule system result in more justice? After all, under a liability rule it is the injurer who pockets the entire efficiency surplus from non-performance ($\phi$), and leaves the victim with expectation damages at best. We have two annotations to make: First, a trade agreement featuring liability-cum-expectation damages objectively is the most efficient contract for self-interested and rational policymakers – given the contractual circumstances, constraints and trade-offs that the bargaining context entails. This makes it a desirable contract from the selfish point of view of the negotiating policymakers. Second, whether the outcome is to be referred as good, fair, equitable, or just, lies in the eye of the beholder. Fairness, equitability or justice are not only inherently subjective, but also relative concepts, and we feel this is not the place to delve into a discussion of comparative justice. If by “good contract” the observer means to imply that the trade agreement should be conducive to the general global welfare of non-signatories – such as consumers and producers worldwide – then we must leave the design of a “good contract” to future research. We can only make an educated guess here: The fact that a contract featuring liability-cum-expectation damages makes the contract more reliable, more incentive-compatible, and induces country governments to be more compliant, cheat less, and make bigger upfront tariff cuts would suggest that this contract at least is not to be judged as right-out pernicious to global welfare.
E. Conclusions and policy recommendations

Our model has shown that the inclusion of a trade policy flexibility mechanism organized as a liability rule backed by expectation remedies Pareto-dominates all other escape designs. Reasonably rational trade policymakers negotiating behind the veil of ignorance can be assumed to opt for this flexibility regime.

- First, compared to other flexibility regimes, a remedial regime organized as LR-cum-expectation damages has two attractive features, namely optimal cooperation and optimal integration. For one, since any future injurer (*he*) will be forced to internalize the externalities he creates, he will be hesitant to enact the escape clause all too often. Pure cooperation (as opposed to the use of escape, which does not yield trade efficiencies) is chosen in those instances, where the CCC would also mandate contractual performance. In addition, expectation damages facilitate far-reaching tariff liberalization at the beginning of the contract. The prospect of increased compliance through the inclusion of the optimal escape scheme into the trade agreement induces both contracting policymakers to consent to their politically optimal level of trade liberalization in the initial negotiations.

An intuitive result of deeper *ex ante* trade liberalization concessions and more instances of pure cooperation in the face of protectionist shocks is that the expected per-period payoff (or continuation value of the game) under an expectation damages policy is strictly higher than under any other trade policy flexibility regime. This finding stands in plain contrast to the contentions of influential IR scholars (e.g. Yarbrough and Yarbrough 1987, Downs and Rocke 1995, Setear 1997, Yarbrough and Yarbrough 1997, Goldstein et al. 2000, Goldstein and Martin 2000, Smith 2000, Rosendorff and Milner 2001, Rosendorff 2005). According to these authors, the inclusion of flexibility instruments into a trade agreement invariably leads to a credibility loss of the system, and to a smaller continuation value than an agreement without escape clauses. This, so the argument continues, is bound to result in less *ex ante* reciprocal market access concessions. We disagree with these: As we have demonstrated, a well-crafted trade policy escape regime with truly commensurate remedies replicates full compliance by the injurer. Receiving the exact replacement value insures the victim (*she*) of an escape measure. We hence should not expect a loss of trust in the system, as measured in terms of the continuation value of per-period cooperation. Rosendorff's formal finding that the inclusion of an opt-out mechanism lowers the value of the agreement for its Members (2005, p. 396) is an outflow of the author's somewhat reductionist modeling – not of the logic of escape clauses.

- Second, as Proposition 2 stated, efficient breach is the *sine qua non* of efficient trade policy flexibility. Advocates of a rigorous introduction of a PR into the WTO are right in stating that renegotiations are apt to lead to efficient breach (confer footnote 22 above): Given that expectation damages are the “compensation floor” for victim countries, *ex post* escape from previously agreed tariff concessions occurs only in those instances where doing so is globally welfare-enhancing. It is important to see that renegotiations *only* lead to efficient breach, if signatories actually reach a

---

95 This finding is confirmed by Herzing (2005, chapter 3, Proposition 5).
96 Ethier (2001), Sykes (1991), and Herzing (2005, chapter 3, Proposition 6 at p. 94) confirm this result.
97 Confer Setear (1997) who argues that trade policy flexibility is a step backward in the process towards greater cooperation, since its relative ease of use increases opportunities for non-cooperation and the likelihood of defection.
mutually agreed solution. A renegotiation breakdown (which happens with a strictly positive probability), however, prevents breach from happening at all. In addition, time costs of lengthy bargains destroy the distributable non-performance gains.

- Third, our results may also disappoint proponents of the introduction of an inalienability rule into the WTO: Prohibiting any sort of *ex post* escape inevitably brings down policymakers’ trust in the WTO Agreement: Under-breach on the part of the injurer and inefficiently high levels of *ex ante* trade liberalization are the results of banning trade policy flexibility.

- Fourth, we have formally shown the pernicious implications of a zero- and a reliance rule of remedies. Doing so struck us as vital, for two reasons. On the one hand, we believe that popular *de facto* WTO escape mechanisms, namely violation of the Agreement, AD and CvD actions, entail exactly one of these two damage rules: If these informal opt-outs go unchallenged by the victim, escaping contractual obligations is for free, tempting injurers to breach inefficiently often. Anticipating this opportunistic behavior, signatories can be expected to respond by scaling down their trade liberalization efforts *ex ante*, or in future trade rounds. Even if the victim challenges these informal opt-outs in front of the WTO dispute settlement body, she is granted reliance damages – at best.98

On the other hand, all *de iure* escape clauses either feature zero damages (e.g. Art. XXII, XV, XVIII GATT), or are under the same *equivalence standard* as is WTO enforcement under the purview of Art. 22.4 DSU. Compensating the victim country with reliance damages, however, implies dynamics similar to the zero-damage measure: The injurer over-breaches, while tariff concessions are inefficiently low.

It is not our concern to explain why the current trade policy flexibility regime in the WTO is a far cry from how we would expect rational trade negotiators to organize contractual escape. Instead, we present the following policy recommendations as a direct outflow of our prior analysis. We have five points to make:

- First, as a bare reform minimum, we would advocate for a novel interpretation of the *equivalence standard* in the WTO. As we have shown, setting *equivalent damages* equal to the reliance remedy is profoundly myopic. Dispute panels should change their damage calculation and adopt the expectation damage rule. This novel interpretation will not only affect trade disputes and WTO enforcement pursuant to Art. 22.4 DSU, but spill over to the enactment of GATT-compatible trade contingency measures.

- Second, we would urge WTO Members to adopt a broad interpretation of GATT Art. XIX. We have shown that an *unconditional* liability rule is the Pareto-dominant trade policy flexibility mechanism (at least when it comes to trade in goods). Art. XIX GATT comes closest to this ideal—if it were not burdened by such a strong enactment threshold and a confining application scope (confer footnote 38 and accompanying text). A lenient interpretation of Art. XIX GATT could be an intermediary solution before the article is in line for a proper redrafting in the long term.99

---

98 WTO arbitrators have repeatedly interpreted the “equivalence” standard to mean direct trade effects. Direct trade effects at most re-establish the *status quo ante* the breach, i.e. are reliance damages in our nomenclature.

99 For concrete redesign suggestions and reformulations of Art. XIX GATT, see for example Roitinger (2004, pp. 194) and Schropp (2008, chapter 8.1).
Third, the WTO membership must make serious efforts to close the loopholes that antidumping and countervailing duties tear into the international trading system. It would be desirable if Members tackled the reform of the AD and CvD codes in a manner that fits the WTO’s mandate and parties’ original intent. It seems that AD and CvD actions today are predominantly used as protectionist opt-out tools, and that the recourse to them as “unfair trade remedies” is a barely veiling fig leaf. Rather, AD and CvD action can be accurately described as a zero-damages liability rule granted to the injurer. This is not the place for delving into a discussion of trade remedy reform.100 We would, however, like to state that we think it best to eventually do away completely with the two codes and to integrate AD and CvD into domestic competition law.101

Fourth, if WTO Members – for political, social, or practicability reasons – wish to attach more importance to the flexibility tool of renegotiations in the future, our policy recommendation is the following: Heeding our results from Appendix A we know that renegotiations are an efficient tool only when mutual agreement is reached, and rapidly so. However, the membership cannot force renegotiating parties to reach a solution – doing so would change the set-up of the game and invite strategic gamesmanship. Thus, coming to a speedy solution should be made as attractive as possible to the bargaining parties. Hence, we would recommend to shorten the renegotiation period considerably (e.g. to one calendar month). After the lapse of this renegotiation period the injurer country should be allowed to escape his obligations. A WTO arbitrator should be mandated to calculate the expectation damages incurred by the victim(s), and to add half of the non-performance gains (or the arbitrator’s best estimate thereof) on top.102

Fifth, our model of trade policy flexibility assumed efficient enforcement. Yet the rules of the game are currently not sufficiently protected against bad-faith extra-contractual behavior, such as when an injuring Member flat-out refuses to pay damages to the victim of a protectionist measure. Violation of the Agreement in the current WTO system can be abused as a de facto escape mechanism sanctioned by reliance damages, nothing deters WTO Members from engaging in extra-contractual behavior – except maybe possible reputation losses (the “name-and-shame factor”). Thus, as soon as efficient breach is safeguarded by above-mentioned reforms, the WTO membership should come up with reforms towards strict and binding enforcement mechanisms. A possible solution may be collective retaliation by the membership (e.g. Maggi 1999), tradable remedies (Bagwell et al. 2005, Limao and Saggi 2006), or withdrawal of certain membership rights of the culprit (e.g. Charnovitz 2001, Lawrence 2003).

100 The reader is referred to the works of Lindsey and Ikenson (2003), Hoekman and Mavroidis (1996) or Bown (2002b). Most of these authors discuss the substitution of AD by antitrust regulation. Horlick and Palmer (2002) focus on the relationship between CvD and antitrust.
101 Antitrust agencies deal with anticompetitive and monopolistic tendencies on a daily basis and thus would appear to be the obvious candidates for assessing unfair practice in international trade (Messerlin 2000, Barfield 2005).
102 The intuition behind adding half of the efficiency surplus from non-performance is simple: Under a PR, victims stand a good chance of appropriating some of the efficiency gains from escape. If, pursuant an infringement of an entitlement protected by a PR, the arbitrator only awarded expectation damages (which grant all the gains from non-performance to the injurer), no injurer would ever choose the route of renegotiation in the first place. Instead, the injuring Member would wait to be sued and receive a higher reservation utility from reimbursing the victim with expectation damages.
Appendix A: Renegotiations with asymmetrical information

There are two countries, victim \((V; \text{she})\) and injurer \((I; \text{he})\). At the beginning of every tariff-setting stage-game, nature reveals the political economy shock \(\theta\) to one of the two countries, which consequently assumes the role of player \(I\) (confer Figure 2 in the main text). We model renegotiations as an alternate request-offer process of three and four periods.\(^{103}\) Both countries bargain over the size of compensation payments that injurer \(I\) has to incur in exchange for victim \(V\)'s waiver of demanding performance as contractually agreed upon. Depending on the size of the unforeseen shock, the requesting country will decide whether to file for renegotiations or not - cognizant of the fact that \(V\) can always threaten to cancel renegotiations and instead insist on the injurer’s specific performance obligation. This, of course, means that rational behavior of the victim will induce her to demand at least expectation remedies as compensation.\(^{104}\)

\[ S(\tau; \theta) \text{ be defined as the excess surplus of breach, i.e. the efficiency gains from non-performance over and above the loss in expectation suffered by the victim.} \]

\[ S = \phi - d_{exp}, \text{ or equivalently} \]

\[ S(\tau; \theta) = W^C(\tau; \theta) - W^V(\tau; \theta) - [V^C(\tau) - V^V(\tau)] . \quad \text{(A1)} \]

Note that \(S\) is a function of both the level of reciprocal tariff liberalization and the unforeseen shock. Consequently, the size of \(S\) is private knowledge to the injurer.

As long as the two signatories have not reached a mutually agreed resolution over how to divide \(S\), the injurer is barred from enacting his desired protectionist policy. The longer renegotiations drag on and the inaction persists, the more the excess surplus of non-performance “melts away” – \(S\) decreases at a steady rate until it completely disappears at some finite date \(T\) (see footnote 63 and accompanying text). Between period \(t_0\) and \(T\) the renegotiating parties have the opportunity to respond to each others’ alternate offers. If the initial offer gets rejected, the other party makes a counteroffer, while \(S\) is “melting” linearly: After \(T/2\), the remaining surplus equals \(S/2\). The number of offer-rounds \(n\) is exogenously predetermined such that \(n \in [1...T]\). According to the assumptions in connection with \(\theta\), \(S\) is drawn from a linear distribution \(F\) with positive density \(f\) over \([0,X]\). \(P_t \in [P_1, P_2, ..., P_T]\) represents the compensation offers as a share of the remaining surplus \(S\) at time \(t \in [I, 2, ..., T]\).

The injurer enters into renegotiations with the clear intention to maximize his share of \(S\). In order to minimize the compensation payable to the victim party, the informed injurer is tempted to misrepresent reality, i.e. to understate the size of the shock, and therewith the actual benefit from non-performance. The victim has all the reasons to mistrust the injurer’s assertions, and uses her power to defer agreement with the aim of forcing the injurer to reveal his private information. \(I\) thus can use

\(^{103}\) We use a three- and a four-period bargaining sequence as representatives of an odd- and even-numbered request-offer game of finite duration, where the victim or the injurer respectively has the right to make the final offer. Assuming more bargaining rounds does not radically alter the presented results.

\(^{104}\) The rational victim will not settle at a loss (and expectation damages put her into a position where she is exactly indifferent between performance and non-performance). Ideally, \(V\) would want to capture all the injurer’s efficiency gains from non-performance for her willingness to “let go” of her contractual rights.
time to convince the victim of the low nature of his suffered shock. For an injurer subject to a high protectionist shock, waiting is too costly, and he is inclined to accept the victim’s initial offer.

A1 Three-period sequential bargaining game

Signatory $V$, the victim country, makes the initial offer for compensation payments.\textsuperscript{105} $V$ only knows that the injurer’s excess surplus from non-performance is linearly distributed in $S=\left[0,X\right]$. $X$ thus constitutes the maximum possible realization of $S$. To maximize her expected payoffs in a three-period game of renegotiation, $V$ must construe a complete plan of action before renegotiations begin. Ignorant of the real shock (and thus of the size of $S$) $V$ rationally distinguishes 4 different “types” of injurers (types A-D),\textsuperscript{106} and anticipates that renegotiations will unfold in the following way: To test the waters, player $V$ starts out setting her request for compensation such that it will only be attractive to the highest type of injurers, i.e. those countries which are under the impact of a high shock $\theta$, and hence dispose of a large distributable $S$.\textsuperscript{107} The initial offer $P_1$ is then rejected by intermediate and low types of $I$. Although injurer types B, C, D possibly suffer from a significant shock, they nevertheless prefer to wait one period and to table a counteroffer. Before renegotiations commence, $V$ already knows the lowest level of $I$’s counteroffer ($P_2$) that she will deem acceptable. $V$ foresees that type-C and type-D injurers will make unacceptable counteroffers in round 2. Only type-C injurers will accept $V$’s last-round offer of $P_3$; type-D injurers will not be able to shoulder compensation payments of $P_3$. Figure A1 illustrates:

![Figure A1: The victim’s decision problem of renegotiations in a three-stage sequential bargain](image)

Ignorant of the true $S$ of the injurer, $V$’s task is to find the optimal threshold levels $x_{A,B,C}$ that maximize her expected welfare, and that determine her decision for $P_1$ and $P_3$, as well as what she considers an acceptable counteroffer $P_2$. $V$’s maximization problem for a three-period bargaining game hence can be written in the following way:

\begin{align*}
\max_{P_1, P_2, P_3} E[U(V)] = & x_A P_1 + x_B P_2 + x_C P_3 + x_D P_4 \\
\text{subject to:} & \quad 0 \leq P_1 \leq x_A, \quad P_2 \leq x_B, \quad P_3 \leq x_C, \quad P_4 \leq x_D,
\end{align*}

\text{105} We follow standard game theory and have the uninformed party make the first offer in order to extract information from the other party.

\text{106} The reason why $V$ distinguishes four types of injurers is an outflow of having three bargaining rounds, or, more precisely, of having a renegotiation process where $V$ can launch two original offers.

\text{107} We assume that the victim uses fixed \textit{ex ante} estimates of $I$’s types and refrains from utilizing Bayesian updating as a bargaining tool out of fear of developing a reputation as a “soft negotiator” in future tariff renegotiations. As a result, $V$ \textit{always} mistrusts $I$’s assertions concerning the size of $S$. More on that in our discussion of outcomes in subsection A.3 infra.
\[
\begin{align*}
\text{Max}_{x_a, x_b, x_c} V &= \frac{X - x_a}{X} P_1 + \frac{x_a - x_b}{X} P_2 + \frac{x_b - x_c}{X} P_3 \\
\text{s.t.} & \\
(1) \quad x_a \geq x_b \geq x_c; \\
(2) \quad x_a - P_1 = \frac{2}{3} x_a - P_2; \\
(3) \quad \frac{2}{3} x_b - P_2 = \frac{1}{3} x_b - P_3; \\
(4) \quad \frac{1}{3} x_c - P_3 = 0.
\end{align*}
\]

\(V\)'s expected return consists of three terms, (i) her initial offer \(P_1\), (ii) \(I\)'s acceptable counteroffer \(P_2\), and (iii) her final offer \(P_3\). All three payoffs are weighted by their respective probabilities. Constraint (2) can be explained by \(I\)'s objective to maximize the difference between \(S\) and the payable compensation \(P_1\) in the first bargaining round: Depending on his true state of \(S\), \(I\)'s first-best option may be to accept \(P_1\) straight away. An injurer with a surplus equal to \(x_a\) is indifferent between accepting the first offer \(P_1\) or making an acceptable counteroffer \(P_2\). The same trade-off goes for constraint (3). For the lowest \(I\)-type it pays off to wait until the final third round. Constraint (4) states that by constructing the final offer, the victim anticipates that any type of \(I\) will accept it as long as \(I\)'s non-performance surplus net of compensation payments is non-negative.

Inserting constraints (1)-(4) into equation (A2), we obtain an unconstrained maximization problem dependant on \(x_a, x_b\) and \(x_c\):

\[
\begin{align*}
\text{Max}_{x_a, x_b, x_c} V &= \frac{1}{3X} \left[X x_a + X x_b + X x_c - x_a^2 - x_b^2 - x_c^2\right].
\end{align*}
\]

(A3).

Constructing the first order conditions with respect to \(x_a, x_b\) and \(x_c\) yields

\[
\frac{\partial V}{\partial x_{A,B,C}} = x_a = x_b = x_c = \frac{X}{2}.
\]

(A4)

The corresponding offers are easily found:

\[
\begin{align*}
P_3 &= \frac{1}{3} x_c = \frac{X}{6}; \\
P_2 &= \frac{1}{3} [x_c + x_b] = \frac{X}{3}; \\
P_1 &= \frac{1}{3} [x_c + x_b + x_a] = \frac{X}{2}.
\end{align*}
\]

(A5)

FOC (A4) reveals a surprising outcome: A game with 3 predetermined negotiation rounds boils down to a “take it or leave it” (TILI) offer by \(V\) in round 1 of the renegotiations. It turns out that it is not in \(V\)’s interest to discriminate \textit{ex ante} between 4 distinct types of injurers – the victim’s \textit{ex ante} priors of all types of injurers are of identical value and equal to \(X/2\). In other words, \(V\) divides the range of possible \(I\)-types into two groups instead of four. Also, the victim is not willing to make price concessions in later bargaining rounds and instead confronts her opponent \(I\) with the same offer of \(X/2\).
(or what is left of it over time) in every round. Hence, the bargaining game reduces to a single-round TILI offer: Given stationarity in the bargaining relationship, no injurer that does not already accept the victim’s initial offer $P_1$ will accept $P_3$ later on in the final stage. Hence, agreement is struck after the first round – or not at all, given that the victim can credibly refrain from updating her beliefs concerning the opponent type after each unsuccessful renegotiation round.

$V$’s optimal action to split the group of possible $I$-types in two instead of four groups reduces the chances of reaching an agreement to 50%. Yet, it also yields welfare-optimal outcomes for $V$. The expected payoff from renegotiation for the victim is given by

$$E(V) = \frac{1}{2} \left[ V + (d + P_1) \right] + \frac{1}{2} V = \frac{1}{2} \left[ V + \left( d + \frac{X}{2} \right) \right] + \frac{1}{2} V^C$$

(A6)

The expected payoff from renegotiations for injurer $I$ equals

$$E(W) = \frac{1}{2} \left[ W - (d + P_1) \right] + \frac{1}{2} W = \frac{1}{2} \left[ W - \left( d - \frac{X}{2} \right) \right] + \frac{1}{2} W^C.$$  

(A7)

We see that engaging into renegotiations yields a strictly higher level of welfare for the victim country than continued cooperation.

Figure A2 illustrates the solution to a three-period sequential bargaining game graphically. An injurer subject to a relatively high shock reaches an agreement in the first renegotiation period; for lower types of $I$ renegotiations will turn out not to be fruitful.

---

Since waiting is costly, the surplus reduces to $2/3 S$ in period 2.
A2 Four-period sequential bargaining game

Next, we describe a bargaining game which lasts for an even number of negotiation rounds, for example four periods. The main difference between a three- and a four-period game is that the injurer gets to make the final offer. Consequently, \( I \) has more bargaining leverage. He can at least secure the surplus that has not “melted” away at the end of penultimate bargaining round (1/4 of the initial surplus \( S \)). \( V \) will accept \( I \)'s offer for \( S \) as long as it is nonnegative.

The victim’s maximization problem looks very similar to equation (A3) above:

\[
\max_{x_a, x_b, x_c} \frac{X - x_a}{X} P_1 + \frac{x_A - x_b}{X} P_2 + \frac{x_b - x_c}{X} P_3
\]

\( s.t. \)

(1) \( x_A \geq x_B \geq x_C \);

(2) \( x_A - P_1 = \frac{3}{4} x_A - P_2 \);

(3) \( \frac{3}{4} x_B - P_2 = \frac{2}{4} x_B - P_3 \);

(4) \( \frac{1}{2} x_C - P_3 = \frac{1}{4} x_c \).

The main difference to the three-period renegotiation game is that the victim’s offer in the penultimate round has less “teeth” in the sense that the injurer can easily reject it and suggest a split which secures \( I \) approximately all remaining surplus and leaves \( V \) with next to nothing. As a result, constraint (4) in equation (A8) is different from that in (A5).

Inserting constraints (1)-(4) into (A8), we again obtain an unconstrained maximization problem dependant on \( x_A, x_B \) and \( x_C \):

\[
\max_{x_a, x_b, x_c} \frac{1}{4X} \left[ X_x_A + X_x_B + X_x_C - x_A^2 - x_B^2 - x_C^2 \right].
\]

(A9)

Constructing the first order conditions with respect to \( x_A, x_B \) and \( x_C \) yields

\[
\frac{\partial V}{\partial x_{A,B,C}} = x_A = x_B = x_C = \frac{X}{2}.
\]

(A10)

The corresponding offers are

\[
\begin{align*}
P_4 &= 0; \\
P_3 &= \frac{1}{4} x_C = \frac{X}{8}; \\
P_2 &= \frac{1}{4} [x_C + x_B] = \frac{X}{4}; \\
P_1 &= \frac{1}{4} [x_C + x_B + x_A] = \frac{3X}{8}.
\end{align*}
\]

(A11)

Just like in a three-period game, the uninformed victim adjusts her initial offer such that agreement is reached with a probability of 50% in the first round. Hence, in an even-numbered sequential bargaining game a mutually agreed solution is either reached in the first round, or in the final round of
renegotiations. High I-types (i.e. those injurers exposed to high shocks) will concede instantaneously to V’s compensation request in round 1, lower I-types will defer agreement until the last round. The expected payoffs from renegotiation for the victim hence equal

\[ E(V) = \frac{1}{2} \left[ r^S(\tau) + \left( d_{\exp} + P_1 \right) \right] + \frac{1}{2} \left[ \frac{3}{4} r^C(\tau) + \frac{1}{4} \left[ r^S(\tau) + d_{\exp} \right] + P_1 \right] \]

\[ = \frac{1}{2} \left[ r^S(\tau) + \left( d_{\exp} + \frac{3X}{8} \right) \right] + \frac{1}{2} \left[ \frac{3}{4} r^C(\tau) + \frac{1}{4} \left[ r^S(\tau) + d_{\exp} \right] \right] \]  \hspace{0.5cm} (A12)

\[ = \frac{5}{8} r^S(\tau) + \frac{5}{8} d_{\exp} + \frac{3}{8} r^C(\tau) + \left( \frac{3X}{16} \right) = r^C(\tau) + \left( \frac{3X}{16} \right), \]

while the expected payoff for the injurer is given by

\[ E(W) = \frac{1}{2} \left[ W^D(\tau, \theta) - \left( d_{\exp} + P_1 \right) \right] + \frac{1}{2} \left[ \frac{3}{4} W^C(\tau, \theta) + \frac{1}{4} \left[ W^D(\tau, \theta) - d_{\exp} \right] + P_1 \right] \]

\[ = \frac{1}{2} \left[ W^D(\tau, \theta) - \left( d_{\exp} + \frac{3X}{8} \right) \right] + \frac{1}{2} \left[ \frac{3}{4} W^C(\tau, \theta) + \frac{1}{4} \left[ W^D(\tau, \theta) - d_{\exp} \right] \right] \]  \hspace{0.5cm} (A13)

\[ = \frac{5}{8} W^D(\tau, \theta) + \frac{3}{8} W^C(\tau, \theta) - \frac{5}{8} d_{\exp} - \left( \frac{3X}{16} \right) . \]

Three observations are in order: First, in a four-period game the expected return from renegotiation for V is smaller than in the three-stage game by 1/16 X, because \((3X/16)<(X/2)\). This is a direct result of the victim’s decreased bargaining power due to I’s possibility to make the last offer. In the same vein, the share of S to be paid by I in return for contractual relief falls from \((X/2)\) to \((3X/16)\). Second and more importantly, the probability of having to perform according to the contract despite severe regret reduces to zero. Under even-numbered games, all renegotiating parties reach an agreement sooner or later. This increases the expected return to the injurer. Third, the expected welfare loss to V is smaller than the increase in welfare for I. So, we can conclude that generally an even-period renegotiation game is globally preferable to a game featuring an odd number of bargaining rounds.

We graphically represent the four-period solution in Figure A3:

---

Figure A3: Outcome of a four-period bargaining game

---

0 \[\rightarrow\] P = 3X/8; P = 0

Agreement is reached in the last negotiation round with probability \(\frac{1}{2}\)

Acceptance of the initial offer with probability \(\frac{1}{2}\)
A3 Interpretation of the results

Triff renegotiations with asymmetrical information bear the following characteristics:

- Renegotiations produce inefficiencies. Not only is bargaining expensive in terms of opportunity-, litigation-, and information searching/gathering costs (not modeled here), but it induces additional waste, because non-performance efficiencies are not captured immediately: In our model, where renegotiations feature an even number of periods, the injurer waits till the very last round to capture what is left of the excess surplus. In odd-numbered games no mutually agreed solution is found half of the time, and consequently the shock-ridden country is prohibited to act upon his regret contingency.

- Independently of which party gets to make the last offer, the victim can always expect to receive some spoils of the non-performance efficiencies on average. Being staffed with a “compensation floor” of at least expectation damages, she is in the comfortable position to “sit out” a lengthy renegotiation process. This hold-out behavior is globally inefficient, but it is rational from the victim’s point of view.

- Independently of which party gets to make the last offer, the injurer has no incentive to reveal the truth: If he disclosed his true level of $\theta$ and therewith of $S$, the victim would always want to appropriate the entire portion of $S$ – and would succeed in doing so, because the costs of inaction are entirely borne by the injurer (cf. Rubinstein 1982).

- In our model, even-numbered bargaining games are globally preferable to games featuring an odd number of periods. Odd-period games bear the risk of negotiation breakdown, which is detrimental to welfare.\(^{109}\)

The bargaining solution we have found is very much driven by our assumption that the uninformed victim can credibly assert that she is not willing to update her beliefs concerning the injurer’s real $S$-type in the course of renegotiations.\(^{110}\) This supposition may seem to contrast state-of-the-art bargaining theory, which tends to integrate Bayesian updating into the negotiation process (cf. the seminal contributions of Harsanyi 1968, Rubinstein 1985, Gul and Sonnenschein 1988). Here, Bayesian updating would imply that pursuant to a rejection of $V$’s initial offer by $I$, the victim infers that $I$ must have been hit by a smaller shock and therefore to be of a lower $S$-type (type-B or less). Updating her belief about $I$’s true nature, $V$ would consequently lower her offers $P_{1,3,5,\ldots}$ in an effort to realize at least some gains from renegotiation, and to minimize the chances of a renegotiation breakdown. Generally, the main difference between “our” renegotiation model and sequential bargains with Bayesian updating is that the price for contractual relief under the latter approach is lower on average.

Bayesian updating may be the strategy of choice in non-repetitive contexts, where renegotiations are a singular event and a breakdown of the bargain is the least desired outcome for both parties, indeed the worst-case scenario. However, we believe that in a context such as ours, Bayesian updating leads the victim into a time-inconsistency trap, and thus to lower payoff expectations in the future: In a trade

\(^{109}\) We caution against generalizing this statement. It seems a direct outflow of our model. However, we believe that the general intuition of this statement equally holds for other modeling setups.
agreement, where renegotiations occur frequently (potentially in every stage-game), the victim country should not develop a reputation as a “soft negotiator”. $V$ therefore needs to credibly assert that she is unwilling to update her beliefs. For if injurer $I$ knows in advance that $V$ is indeed a Bayesian updater and therewith inclined to steadily lower her request for compensation with every rejection by $I$, “higher” injurer-types will have an incentive to misrepresent their identity so as to settle for a more lenient offer later on. $V$’s announcement offer $P_I$ then is not credible in light of $I$’s anticipation of later leniency. In a game without Bayesian updating the victim’s outcome of the stage-game is significantly lower.

To avoid this kind of time-inconsistency dynamics, $V$’s best strategy is to pre-commit to a tough stance in renegotiations and to stick with the compensation offer she once made. In so doing, she may risk breakdown of renegotiations in some stage-games, but in the long run this strategy pays off: Assuming that each country will undergo many renegotiations in the future, $V$ appreciates a higher average per-period payoff. Every victim will thus make clear from the outset that she is not inclined to buy into $I$’s assertions concerning his type. To the dominant victim strategy is to be a tough negotiator and to refrain from Bayesian updating, not least having the reputation as a tough negotiator induces more injuring countries to agree in the first round of renegotiations.

110 As we stated above, $V$’s prior of different $I$-types is stable and equals $X/2$ in any given negotiation round.

111 Technically, the distinction between a Bayesian updating game and a commitment game of the kind discussed here is that the former is solved by backwards induction, whereas the latter is characterized by finding an ex ante (predetermined) strategy which yields the highest expected return.

112 Common to all asymmetrical information renegotiation games is the feature that bargaining generally lasts more than one period, and that welfare is lost due to lengthy procedures. Hence, even in a sequential bargaining game with Bayesian updating it may be possible that renegotiations lead to significant inefficiencies (as compared to the CCC). In a request-offer game with Bayesian updating, the injurer uses “waiting” as a signal to convince $V$ of his low type. Yet signals are costly. Hence, the Bayesian variant of renegotiations may ultimately entail inefficiencies, since agreement is deferred voluntarily by the injurer. These inefficiencies, in turn, are likely to have a negative effect on injurers’ breach behavior later on, and the reciprocal ex ante trade liberalization behind the veil of ignorance.
Bibliography


