International Law and Justice Working Papers

IILJ Working Paper 2013/5

NON-VIOLATIONS

ROBERT STAINGER AND ALAN SYKES

Institute for International Law and Justice
139 MacDougal Street, 3rd Floor
New York, NY 10012
Website: www.iilj.org

ISSN: 1552-6275 © Robert Staiger and Alan Sykes

Working Papers are issued at the responsibility of their authors, and do not reflect views of NYU, the IILJ, or associated personnel.

New York University School of Law
New York, NY 10012
U.S.A.

Cite as:
IILJ Working Paper 2013/5
(www.iilj.org)

Non-Violations*

Robert W. Staiger
Wisconsin and NBER

Alan O. Sykes
New York University Law School

July 2013

Abstract

The "non-violation nullification or impairment" doctrine is among the more unique and perplexing features of WTO law. We examine the scope and application of the non-violation doctrine in the GATT/WTO. Observed cases raising non-violation claims are relatively uncommon, the non-violation claims that are observed are usually not adjudicated, and those that are adjudicated are unsuccessful in a relatively high percentage of cases. This paper offers a number of reasons for the relatively modest (observed) role of the non-violation doctrine over the history of the WTO/GATT system. We report the results of a formal model developed in a companion paper, which can deliver differences in equilibrium usage of violation and non-violation claims that mirror these broad features. We note that although non-violation claims are not often observed and when observed are less successful, they can still serve a valuable role in the GATT/WTO, owing in part to their "off equilibrium" effects. We develop a range of other considerations informally, which further help to explain the limited role of non-violation claims in the system and hint at some possible reforms to clarify its proper scope.

*We thank Petros C. Mavroidis and seminar participants at the 2012 American Law and Economics Association meetings and the University of Virgina International Law Workshop for helpful comments.

1. Introduction

The "non-violation nullification or impairment" doctrine is among the more unique and perplexing features of WTO law. It is contained in Article XXIII of the original 1947 version of GATT, which provides in pertinent part:

> If any contracting party should consider that any benefit accruing to it directly or indirectly under this Agreement is being nullified or impaired or that the attainment of any objective of the Agreement is being impeded as the result of
>
> (a) the failure of another contracting party to carry out its obligations under this Agreement, or
>
> (b) the application by another contracting party of any measure, whether or not it conflicts with the provisions of this Agreement, or
>
> (c) the existence of any other situation,
>
> the contracting party may [have recourse to the dispute resolution process]...

On its face, Article XXIII creates six distinct "causes of action:" three for "nullification or impairment" of benefits, and three for situations where the "attainment of any objective... is being impeded." Within each group, an action may be brought for "failure... to carry out... obligations," for the "application by another... of any measure, whether or not it conflicts" with GATT, or "any other situation."

In GATT practice, and since the creation of the WTO in 1994 (which subsumes GATT), only two of these actions have proven significant. Those associated with impeding the "attainment of any objective" have essentially disappeared, as have the actions concerning "any other situation." The actions that arise in modern practice all focus on "nullification or impairment." The so-called "violation" cases invoke Article XXIII(1)(a), while "non-violation" cases proceed with respect to Article XXIII(1)(b). Non-violation claims are the focus of this paper.

Although the history of non-violation claims under GATT has been documented elsewhere, and economists have pointed out the useful role that non-violation claims could play in principle, no study has examined the scope and application of the non-violation doctrine in GATT/WTO practice from an economic perspective. As we detail below, the doctrine has proven quite limited in the situations to which it applies. It plays a much smaller role under GATT than the drafters seem to have envisioned, and much less of a role than economic theory might seem

to suggest it should. Among the stylized facts that we seek to explore are certain well-known differences in the observed usage of violation and non-violation claims. In particular, observed filings of non-violation claims have occurred only selectively in comparison to violation claims. Non-violation claims that are filed often do not get adjudicated. Finally, the success rate of the small numbers of non-violation claims that *are* adjudicated is low compared to that of violation claims.

The analysis below will analyze the apparent mismatch between the role of non-violation claims in theory and in practice, and will address more generally the question of what GATT-legal "measures" (or "other situations") should be deemed actionable. The goals of the analysis are both positive and normative – positive in an effort to understand why the non-violation doctrine has played such a limited role in GATT/WTO jurisprudence, and normative in an effort to determine whether a broader role for the doctrine, in line with that envisioned by the drafters and suggested by economic theory, might be desirable. The discussion will also consider whether the case law on non-violation claims provides a coherent account of its own scope and application, and whether the limited use of non-violation claims in GATT/WTO practice presents a puzzle for economic theory.

In a companion technical paper (Staiger and Sykes, 2013), we present a model that allows for both violation and non-violation claims to arise, and show that under plausible circumstances (plausible parameter values in the model) non-violation claims will occur only selectively, will often not be adjudicated when they are filed, and will have a lower success rate than violation claims when they are adjudicated. We sketch these results below. We thus suggest that economic theory can account for the limited use of non-violation claims in GATT/WTO practice. But this raises a further question: Given the small numbers of successful non-violation claims that have been observed, can the non-violation clause serve any valuable role in the GATT/WTO? The answer is yes – even if the observed use of the non-violation claim is quite limited, the ability to bring non-violation claims can still serve a valuable role in the GATT/WTO "off the equilibrium path," a notion that we will explain in detail.

The formal model in our companion paper abstracts from several broader issues associated with non-violation claims and we discuss a number of these here as well. Our discussion suggests that the modest observed role played by the non-violation doctrine over the history of the GATT/WTO system may well be appropriate, and that both the drafters of GATT and inferences drawn directly from existing economic theory may have been overoptimistic about

what the doctrine can realistically accomplish. We suggest some possible modifications to legal doctrine to clarify the limited role that non-violation claims can usefully play.

Section 2 reviews the GATT negotiating history and the subsequent case law on non-violation claims in the GATT and WTO. It then summarizes the state of the legal doctrine and notes some attendant puzzles. Section 3 reviews the existing economic theory on the doctrine, including an introduction to the standard economic theory of trade agreements, while section 4 presents our informal exposition of the results in our companion paper. Section 5 then develops some additional economic considerations that counsel a narrower role for non-violation claims than the existing commentary suggests. Finally, section 6 offers a brief conclusion.

2. The Non-Violation Doctrine in Practice

As Petersmann (1997) reports, the concept of nullification or impairment originated in various bilateral trade treaties that preceded GATT. These earlier treaties also typically allowed for the possibility that nullification or impairment might arise for reasons that did not involve violations of treaty commitments. But the drafters of GATT did not borrow the concept without reflection. Some insight about the drafters' concerns may be gleaned from contemporaneous negotiations over the creation of the charter for the International Trade Organization (ITO), an institution that was to replace GATT but never came into being. Extensive discussions relating to the draft ITO charter centered on the possibility that the benefits of a trade agreement might not be realized for a variety of reasons. Among the prominent concerns was the possibility that the U.S. economy would return to the depression conditions that preceded World War II. Such concerns motivated not only the provision that nullification or impairment might result from "measures" legal under the treaty, but also the possibility that it might result from some "other situation."

The GATT negotiators were mindful of the fact that this language is exceedingly vague and open-ended as to the circumstances that might trigger a right to dispute resolution, and some negotiators were critical of it on that basis. The negotiators nevertheless chose to retain the "any measure" and "other situation" provisions, with the hope that the GATT membership would interpret them sensibly over time. See Hudec (1990), chapter 4.

One negotiator summed up the logic of the approach as follows:

> We shall achieve...if our negotiations are successful, a careful balance of the

> interests of the contracting parties. This balance rests upon certain assumptions as to the character of the underlying situation in the years to come. And it involves a mutuality of obligations and benefits. If, with the passage of time, the underlying situation should change or the benefits accorded any contracting party should be impaired, the balance would be destroyed. It is the purpose of Article XXIII to restore this balance by providing for a compensatory adjustment in the obligations which the contracting party has assumed. [quoted in Petersmann (1997), p. 145.]

The drafters thus conceived of Article XXIII as a broad and flexible mechanism to ensure that the balance of negotiated concessions would not be upset by unforeseen circumstances. Should such circumstances arise for any reason, Article XXIII afforded aggrieved parties the opportunity to negotiate for compensatory adjustments, to withdraw their own concessions if negotiations failed, and even to withdraw from the treaty.

Throughout GATT and now WTO practice, however, formal complaints regarding GATT-legal "measures" and "other situations" have played at best a minor role. No "other situation" complaints succeeded under GATT (only two formal complaints were raised), and none have been brought under the WTO to our knowledge. Complaints about GATT-legal "measures" have also been few in number, as shall be seen below, and only three have resulted in decisions favorable to complainants that were accepted by the GATT membership as binding law ("adopted decisions" in GATT parlance), the last such case arising over two decades ago.

The remainder of this section will review the evolution of non-violation doctrine. Section 2.1 offers a chronology of disputes to date and then summarizes the key principles and implications of the doctrine. Section 2.2 considers whether the existing body of doctrine affords a satisfactory account of the issues addressed by non-violation claims and their resolution.

2.1. The History of Non-violation Complaints

It is not possible to supply a definitive history of all the circumstances in which non-violation claims have been raised because many disputes over the history of the GATT/WTO system have been settled on an informal basis, with little or no record of the precise legal claims at issue. We do know that some cases raised non-violation claims that were never adjudicated because a settlement was reached, some paired violation and non-violation claims, the latter never adjudicated because the violation claims were successful, and some raised inchoate non-violation claims that were not adjudicated because they lacked sufficient particularity. This

section focuses on the non-violation claims that proceeded to the point of a definitive ruling on the merits by a dispute panel.

2.1.1. The GATT Years

All of the successful non-violation claims arose prior to the creation of the WTO. Their facts and key doctrinal points are detailed below. The discussion includes panel reports that were never adopted by the GATT membership.

Australia – Subsidy on Ammonium Sulphate. In 1947, Chile secured a tariff concession on sodium nitrate fertilizer from Australia. At the time of the negotiations, Australia granted subsidies for the purchase of both sodium nitrate fertilizer and ammonium sulphate fertilizer, which was domestically produced. Some time after the conclusion of negotiations, Australia discontinued the subsidy on sodium nitrate only, predictably leading to a shift in purchases toward ammonium sulphate. Australia was under no legal obligation to subsidize sodium nitrate, but the change in treatment was found by a GATT working party (a precursor to the GATT dispute panel process) to constitute nullification or impairment "if the action of the Australian Government which resulted in upsetting the competitive relationship... could not reasonably have been anticipated by the Chilean Government, taking into account all pertinent circumstances and the provisions of the General Agreement, at the time it negotiated the duty-free binding on sodium nitrate." The report emphasized that the fertilizer subsides were originally introduced at the same time, and had always been equal in amount up to the time of the negotiations.

Interestingly, the same working party suggested that it might be a different case if Australia had introduced a new, GATT-legal subsidy on only one of the products. It observed that, "given the freedom under [GATT]... to impose subsidies and to select the products on which a subsidy would be granted, it would be more difficult to say that the Chilean Government had reasonably relied on the continuation of the same treatment for the two products." In effect, the working party hinted that if a member acts pursuant to explicit authority contained in GATT (such as the authority in Article III(8)(b) to introduce a new production subsidy), other parties should not be heard to complain.

A later working party report, adopted during the 1954-55 GATT review session, took a somewhat different view. It concluded that "a contracting party which has negotiated a [tariff]

concession under Article II may be assumed... to have a reasonable expectation, failing evidence to the contrary, that the value of the concession will not be nullified or impaired... by the subsequent introduction or increase of a domestic subsidy on the product concerned." The same working party noted that members were free to negotiate over subsidies and other matters affecting the value of tariff commitments, and to memorialize their understandings in the pertinent schedules of tariff bindings.

Germany – Treatment of Imports of Sardines. Norway secured a tariff concession from Germany in 1951 on its exports of sprats and herring. Germany subsequently afforded more favorable tariff treatment to exports of a competing species of sardine, exported mainly by Portugal. Norway complained, and a dispute panel report adopted in 1952 found in Norway's favor. The panel was unable to accept Norway's argument that the three species were "like products," and thus declined to rule that the tariff discrimination violated the most-favored nation obligation of GATT Article I. Instead, the panel found that the change in tariff treatment "could not reasonably have been anticipated by the Norwegian Government at the time it negotiated for tariff concessions," thus ruling for Norway on non-violation grounds. It emphasized that the three species were closely related and directly competitive with each other, and that they had been treated equally by Germany for tariff purposes between 1925 and the time of the 1951 negotiations.

Germany – Import Duties on Starch. Benelux countries complained in 1954 that Germany had failed to reduce its duties on starch products in accordance with a letter sent by the chief German tariff negotiator in 1951, which promised to negotiate promptly for a reduction in those duties to 15%. The dispute panel ruled that the commitments in the letter constituted part of the balance of concessions in the 1951 negotiations, and that they appeared to contemplate reductions in the German tariff without any further reciprocal concessions by the Benelux nations. The report was never adopted because the case ultimately settled.

EEC – Production Aids on Canned Fruits and Dried Grapes. The United States brought a complaint in 1982 claiming nullification or impairment of benefits relating to certain tariff concessions on fruit products negotiated in 1962, 1967, 1973 and 1979. The basis for the complaint was the introduction of new production subsidies for domestic production of the goods in question. The EEC justified the subsidies as necessary to compensate producers

for the elevated prices of the raw fruits caused by the Common Agricultural Policy (CAP). The United States responded, among other things, that the subsidies were excessive in that respect, more than compensating for such price differentials. The panel held that, at least to the extent that the subsidies overcompensated for the price elevation caused by the CAP, the new subsidies introduced after the tariff negotiations constituted nullification or impairment. The panel stated that the issue was whether they "could not have been reasonably anticipated" by the U.S. tariff negotiators, and the panel ruled for the United States on that issue. The panel report was issued in 1985, but was not adopted because the case settled subsequently.

EEC – Tariff Treatment on Citrus Products. Also in 1982, the United States brought a complaint relating to EEC tariff preferences on citrus products from certain Mediterranean basin countries. The preferences were granted ostensibly in conformity with GATT Article XXIV, which allows for the formation of customs unions and free trade areas. But the United States (and others) complained that the arrangements were not in conformity with Article XXIV. That issue was not conclusively resolved because GATT procedures for the review of customs unions and free trade areas (and interim agreements leading to them) were ineffective, and rarely led to any definitive conclusions. GATT panels refused to intervene, insisting that the (ineffective) review process under Article XXIV(7) was the only review option. Because a ruling under Article XXIV was not obtainable from the panel, the United States argued instead that the preferences represented a non-violation nullification or impairment. The panel recognized some clear difficulties with this argument, in that a number of the tariffs in question were not bound (had not been the subject of tariff concessions by the EEC), while some of the bound tariffs had been negotiated after the Mediterranean preferences were in place, so that the U.S. negotiators knew about them. Nevertheless, the panel ruled for the United States in February 1985, holding that benefits that the United States could reasonably have expected in connection with GATT Article I (the most-favored-nation obligation) had been impaired.

The decision was highly controversial, in that it raised the possibility of widespread nullification or impairment within the system due to numerous preferential arrangements, the legality of which were never definitively adjudicated. It also broke new ground by suggesting that nullification or impairment could be found even in the absence of a negotiated tariff concession, or even if the "impairment" was known to the tariff negotiators at the time of the relevant tariff negotiations. The adoption of the report was blocked and eventually the matter was settled.

EEC – Payments and Subsidies to Processors and Producers of Oilseeds. The United States secured tariff concessions from the EC on oilseeds in 1962. In 1988, some 26 years later, the United States brought a complaint regarding some new EEC subsidy programs that were designed to insulate oilseed producers from competitive harm due to import price fluctuations. Part of the case involved alleged violations of GATT Article III, but it also included a non-violation claim. The panel ruled for the United States on the latter claim, stating that "the United States may be assumed not to have anticipated the introduction of subsidies which protect Community producers of oilseeds completely from the movement of prices for imports and thereby prevent tariff concessions from having any impact on the competitive relationship between domestic and imported oilseeds."

2.1.2. Non-Violation Claims in the WTO

The Uruguay Round of GATT negotiations, with its detailed review of the dispute settlement system, afforded GATT members an opportunity to revisit the role of non-violation complaints. But the negotiators for the most part preserved the status quo ante.

WTO Treaty Text Article 26 of the WTO Dispute Settlement Understanding (DSU) preserves the non-violation complaint and makes clear that the rules and procedures of the DSU apply generally to those complaints (subject to a few caveats), including the rules that allow an aggrieved member to retaliate by suspending its own concessions if a negotiated settlement cannot be reached. The caveats include two principles that emerged in previous GATT practice. Article 26(1)(a) provides that non-violation claims require a "detailed justification." Article 26(1)(b) affirms the fact that if a measure does not violate WTO treaty commitments, but nevertheless nullifies or impairs benefits, there is no obligation to withdraw the measure. Rather, the member in question should make a "mutually satisfactory adjustment." Article 26(1)(c) provides that an arbitration over the reasonable time for a member to comply with a ruling can, in a non-violation case, also include arbitration over the level of nullification or impairment suffered (which would serve as a benchmark for the permissible level of any retaliation). Article 26(1)(d) stipulates that an agreement to compensate the complaining member can constitute a "final settlement" of the dispute.

The "any measure" non-violation claim was also incorporated into the General Agreement on Trade in Services (GATS), Article XXIII(3), with language similar to GATT Article XXIII.

Interestingly, the GATS drafters did not provide for "other situation" complaints. The role of non-violation claims in intellectual property disputes raises more complex issues and has been the subject of some ongoing negotiations, the details of which are unimportant here.

Finally, the WTO Agreement on Subsidies and Countervailing Measures (SCMs) essentially codifies the rule that new subsidies introduced after a tariff negotiation can constitute nullification or impairment of tariff bindings. Such subsidies are subject to the remedial provisions of SCMs Article 7, which require a member to eliminate the adverse effects of the subsidy or to withdraw it, an obligation that goes beyond what is traditionally required in non-violation cases (mutually satisfactory adjustment).

WTO Case Law A number of WTO disputes have included non-violation claims. To date, however, only three cases have proceeded to the point of a panel report that adjudicates the non-violation claim on the merits, and none of the non-violation claims have succeeded.

Japan – Film. In 1996, the United States brought a complaint regarding a wide range of "measures" taken by the Japanese Government that ostensibly had the effect of preventing U.S. manufacturers of photographic film (principally Kodak) from competing successfully in the Japanese market. The measures included, inter alia, various rules regarding foreign investment, the "large stores" law, and certain policies of the Japan Fair Trade Commission and MITI. The panel report, which was not appealed and was adopted in 1998, ruled against the United States on both the violation and non-violation claims. The analysis of the non-violation claims is lengthy and complex, but for the most part the panel ruled that the United States did not carry its burden of proving that the "measures" at issue were causally responsible for the inability of U.S. exporters to penetrate the Japanese market more successfully.

The decision includes perhaps the most detailed exposition of the non-violation doctrine to date. The panel initially observed that non-violation cases have played only a limited role in the history of the GATT system, and proceeded to remark:

> This suggests that both the GATT contracting parties and WTO Members have approached this remedy with caution and, indeed, have treated it as an exceptional instrument of dispute settlement. We note in this regard that both the European Communities and the United States in the EEC — Oilseeds case, and the two parties in this case, have confirmed that the non-violation nullification or impairment

> remedy should be approached with caution and treated as an exceptional concept. The reason for this caution is straightforward. Members negotiate the rules that they agree to follow and only exceptionally would expect to be challenged for actions not in contravention of those rules.

It then sets out a three-part structure for the complainant in a non-violation claim – (i) the application of a "measure;" (ii) the identification of a "benefit" owing to the complainant under some WTO agreement; and (iii) a demonstration that the measure has nullified or impaired that benefit. The complainant of course has the burden of proof.

One issue in the case concerned the question whether market access expectations associated with earlier rounds of tariff concessions remain protected even after new tariff concessions on the same product during a more recent negotiating round. In other words, does the latest set of negotiations on a product extinguish any market access expectations relating to that product that may have arisen during prior negotiating rounds? Citing the GATT Oilseeds panel, the Film panel held that expectations from prior rounds remain protected despite subsequent negotiations or concessions.

To aid in identifying the protected "expectations," the panel introduced a rebuttable presumption that measures already in place at the time of the tariff negotiations at issue are anticipated by the negotiators (and hence do not upset market access expectations). Conversely, measures introduced after the negotiations at issue are presumed not to have been anticipated.

The panel also embraced a principle found in many GATT cases to the effect that nullification or impairment arises when a measure upsets the competitive balance between the complainant's exports and competing goods:

> [I]t must be demonstrated that the competitive position of the imported products subject to and benefitting from a relevant market access (tariff) concession is being upset by ('nullified or impaired ... as the result of') the application of a measure not reasonably anticipated. The equation of 'nullification or impairment' with 'upsetting the competitive relationship' established between domestic and imported products as a result of tariff concessions has been consistently used by GATT panels examining non-violation complaints. For example, the EEC — Oilseeds panel, in describing its findings, stated that it had 'found ... that the subsidies concerned

had impaired the tariff concession because they upset the competitive relationship between domestic and imported oilseeds, not because of any effect on trade flows'. The same language was used in the Australian Subsidy and Germany — Sardines cases.

Thus, the complainant need not prove an adverse impact on trade flows (presumably, a decline in its export volume or market share), only that the competitive balance has been altered.

EC – Asbestos. In 1998, Canada brought a complaint relating to a 1996 French prohibition on the sale (and importation) of asbestos and asbestos-containing goods. The ban was enacted as a health measure. Canada's interest lay mainly in its desire to continue exporting chrysotile asbestos fibers for use in construction applications, such as for the reinforcement of certain concrete construction forms. These uses of chrysotile asbestos, and the sale of the resultant asbestos-containing products, were considered safe by Canada (and other countries) with proper precautions. Canada had obtained tariff concessions on asbestos and asbestos-containing products at various points in time, including 1947, 1962 and the Uruguay Round. The Canadian complaint involved a mix of violation and non-violation claims.

The panel rejected the violation claims, as did the Appellate Body employing somewhat different reasoning. As to the non-violation claims, both the panel and the Appellate Body reiterated the notion in Japan – Film that non-violation claims should be "approached with caution and treated as an exceptional instrument of dispute settlement." But the panel and the Appellate Body rejected the European argument that non-violation claims are limited to matters not addressed specifically elsewhere in GATT. Even a measure that is permissible under the GATT Article XX(b) exception for protection of human health, for example, can be the subject of a non-violation claim. Likewise, both the panel and the Appellate Body rejected the European argument that non-violation claims are limited to "commercial measures" such as subsidies or tariff discrimination, noting that Article XXIII(1)(b) makes reference to "any measure."

The fact that the asbestos ban was enacted for health reasons, however, was not irrelevant – it allowed France to avoid the presumption established in Japan – Film that measures introduced after a tariff negotiation could not reasonably be anticipated. In a portion of the panel report that was not appealed, the panel suggested that measures falling under Article XX (including

public health measures) are subject to a "stricter burden of proof:"

> [W]e consider that the special situation of measures justified under Article XX, insofar as they concern non-commercial interests whose importance has been recognized a priori by Members, requires special treatment. By creating the right to invoke exceptions in certain circumstances, Members have recognized a priori the possibility that the benefits they derive from certain concessions may eventually be nullified or impaired at some future time for reasons recognized as being of overriding importance. This situation is different from that in which a Member takes a measure of a commercial or economic nature such as, for example, a subsidy or a decision organizing a sector of its economy, from which it expects a purely economic benefit. In this latter case, the measure remains within the field of international trade. Moreover, the nature and importance of certain measures falling under Article XX can also justify their being taken at any time, which militates in favour of a stricter treatment of actions brought against them on the basis of Article XXIII:1(b).

It later indicated that the greater burden of proof justified rejection of the presumption in Japan – Film:

> [W]e consider that in view of the time that elapsed between those concessions and the adoption of the [ban on asbestos] (between 50 and 35 years), Canada could not assume that, over such a long period, there would not be advances in medical knowledge with the risk that one day a product would be banned on health grounds. For this reason, too, we also consider that the presumption applied in Japan – Film cannot be applied to the concessions granted in 1947 and 1962. Any other interpretation would extend the scope of the concept of non-violation nullification well beyond that envisaged by the Panel in Japan – Film. On the contrary, it is for Canada to present detailed evidence showing why it could legitimately expect the 1947 and 1962 concessions not to be affected and could not reasonably anticipate that France might adopt measures restricting the use of all asbestos products 50 and 35 years, respectively, after the negotiation of the concessions concerned. In the present case, the burden of proof must be all the heavier inasmuch as the intervening period has been so long. Indeed, it is very difficult to anticipate what a Member

> will do in 50 years time. It would therefore be easy for a Member to establish that he could not reasonably anticipate the adoption of a measure if the burden of proof were not made heavier.

The panel went on to rule that Canada had not met its burden of proof.

Finally, with regard to Canada's expectations arising from the Uruguay Round negotiations, the panel ruled that medical evidence regarding the carcinogenic properties of asbestos was extensive by that time. Accordingly, Canada could not have reasonably expected that WTO members would not at some point in the future act to ban asbestos products.

Korea – Government Procurement. The WTO Government Procurement Agreement (GPA) is a plurilateral arrangement (not all WTO members belong) under which signatories commit to open government procurement to foreign sellers of goods and services. The procurement activities governed by the GPA are identified by the government agency doing the procurement ("entities covered" under GPA Article I). As an example, the United States might make procurement by the Department of Transportation subject to GPA coverage, but decline to make procurement by the Department of Defense subject to coverage. Once an entity is listed as "covered," a number of additional obligations apply.

Korea constructed a new airport at Inchon near Seoul, and imposed a number of restrictions on the participation of foreign suppliers that seemingly clashed with obligations in the GPA, if the Korean entity in charge of procurement for the airport was a covered entity. The United States brought a complaint alleging that it was, but the dispute panel concluded that it was not (the panel report was not appealed).

Article XXII of the GPA, concerning dispute settlement, contains language similar to GATT Article XXIII, and in particular recognizes the possibility of nullification or impairment due to "any measure." The United States thus argued in the alternative that Korea's practices amounted to non-violation nullification or impairment. The argument was difficult to make on the basis of past non-violation precedents, since the panel ruled that the Korean agency in charge of the airport project was not a covered entity – thus, there was no negotiated commitment comparable to a negotiated tariff binding under GATT that could have created expectations of market access for the United States.

The United States argued nevertheless that it had been misled by Korea during the GPA negotiations as to the scope of Korea's commitments and, in particular, as to the entities that

would be responsible for procurement in the airport project. The panel was willing to accept such a claim in principle, and concurred with the United States that Korea had given incomplete answers to U.S. questions during the negotiating process. It nevertheless ruled that the United States had failed to prove its non-violation claim because the Korean legislation that placed the airport project under the jurisdiction of a non-covered entity was in place two years prior to the close of the negotiations, and the United States was charged with notice of that legislation.

The panel's apparent willingness to consider the U.S. claim despite the absence of a negotiated commitment covering the airport project flowed in part from the panel's belief that non-violation claims should be interpreted against the backdrop of customary international law and, in particular, the principle that treaty commitments must be performed in good faith:

> In our view, the non-violation remedy as it has developed in GATT/WTO jurisprudence should not be viewed in isolation from general principles of customary international law. As noted above, the basic premise is that Members should not take actions, even those consistent with the letter of the treaty, which might serve to undermine the reasonable expectations of negotiating partners. This has traditionally arisen in the context of actions which might undermine the value of negotiated tariff concessions. In our view, this is a further development of the principle of pacta sunt servanda... The principle of pacta sunt servanda is expressed in Article 26 of the Vienna Convention in the following manner: "Every treaty in force is binding upon the parties to it and must be performed by them in good faith"...
>
> The non-violation doctrine goes further than just respect for the object and purpose of the treaty as expressed in its terminology. One must respect actual provisions (i.e., concessions) as far as their material effect on competitive opportunities is concerned. It is an extension of the good faith requirement in this sense.
>
> The vast majority of actions taken by Members which are consistent with the letter of their treaty will also be consistent with the spirit. However, upon occasion, it may be the case that some actions, while permissible under one set of rules (e.g., the Agreement on Subsidies and Countervailing Measures is a commonly referenced example of rules in this regard), are not consistent with the spirit of other commitments such as those in negotiated Schedules. That is, such actions deny the competitive opportunities which are the reasonably expected effect of such commitments.

These passages offer an interpretation of the non-violation doctrine not found in prior dispute reports under GATT or the WTO. The linkage to the general "good faith" obligation under customary international law, and the focus on measures that violate the "spirit" if not the "letter" of treaty commitments, hint that the function of non-violation claims is to address deceit and other forms of opportunistic behavior.

After advancing this perspective, however, the panel noted that a non-violation complaint does not require evidence of improper behavior:

> [W]e do not mean to introduce here a new requirement that a complainant affirmatively prove actual bad faith on the part of another Member. It is fairly clear from the history of disputes prior to the conclusion of the Uruguay Round that such a requirement was never established and there is no evidence in the current treaty text that such a requirement was newly imposed. Rather, the affirmative proof should be that measures have been taken that frustrate the object and purpose of the treaty and the reasonably expected benefits that flow therefrom.

2.1.3. Summary

The lessons of the case law to date may be summarized as follows:

- The elements of a non-violation claim are three: (i) the identification of a "measure:" (ii) the identification of a "benefit" under an agreement; (iii) proof that the measure nullifies or impairs the benefit.
- Successful non-violation claims have all involved "commercial measures" such as subsidies and tariffs that change in a way that reduces export opportunities for the complainant. But the doctrine in principle permits complaints about any type of measure (and even " other situations").
- Historically, the paradigm non-violation case was a new (post-negotiation) subsidy to domestic firms that compete with imports. Such measures are now "actionable subsidies" under the SCMs Agreement so that the non-violation claim is no longer needed.
- With two outliers (EC – Citrus and Korea – Government Procurement), the cases concern the protection of expectations associated with specifically negotiated concessions.
- The protected expectations must be reasonable, in the sense that the negotiators must not have been able to anticipate the policy change that nullifies or impairs the concession. A

presumption now exists that post-negotiation measures are unanticipated but that presumption need not apply to health measures and perhaps others of particular importance.

- The protection for expectations does not expire, and neither is it extinguished by subsequent negotiations relating to the same concession.

- Although the non-violation doctrine bears some relation to the familiar principle of customary law that requires good faith in the performance of treaties, it does not require a showing of bad faith and indeed in most of the cases there is no evidence that negotiators acted in bad faith.

2.2. Some Puzzles and Inconsistencies

The case law presents a potentially misleading picture of the non-violation doctrine for reasons familiar to students of the economics of litigation. Cases that litigate to conclusion may not be representative of a potentially broader range of cases that often settle, and some types of disputes may not arise at all because WTO members, in the shadow of the doctrine, may eschew certain types of behavior that would give rise to a valid claim. One must therefore be careful in drawing inferences about the scope and significance of the doctrine from the litigated cases.

Nevertheless, it seems clear that the doctrine has played a smaller role than the original GATT negotiators imagined it might play. In the negotiating history, the drafters imagined that a wide range of unanticipated measures, and even mere "situations," could upset market access expectations and create an entitlement to compensation or withdrawal of comparable concessions. The possibilities included various macroeconomic policies and shocks, regulatory policies, and unexpectedly cheap competition in export markets.

Yet, the set of measures that have been the subject of non-violation complaints, let alone successful ones, has been far more limited. The reported non-violation claims almost always involve "commercial matters" such as subsidies, tariffs, tariff discrimination, and quantitative restrictions. Only a few unsuccessful cases, particularly involving Japan, have involved a broader range of issues such as competition policies, and only EC – Asbestos joined issue with any form of prudential or safety regulation. Macroeconomic policy measures have not been a subject of complaint (not even China's much-maligned currency policies, at least so far), nor have all sorts of regulatory issues that have surfaced in violation cases. Moreover, although non-violation claims could be advanced in any dispute involving an export subject to a negotiated concession, at least as a backup argument, GATT/WTO members advance them only

selectively.

What accounts for the limited role of the non-violation claim? It is certainly not compelled by the treaty text. It is also difficult to find tight limits on non-violation claims in the case law. On its face, the doctrine applies whenever the value of a negotiated concession has been impaired by "any measure" that "could not reasonably have been anticipated by the negotiators." The protection for market access expectations does not expire with time. True, a "detailed justification" is required, but why does that pose a problem? Can complainants not simply point to the state of affairs at the time of negotiations, and to the fact that some recent measure, which was not readily foreseen ex ante, has significantly reduced market access relative to that state of affairs? Although the cases refer to the non-violation remedy as "exceptional," it is hardly obvious why that should be true, and little in the case law offers much help or serves to identify the "exceptional cases" that are appropriate.

One case that reins in the doctrine somewhat, albeit in questionable fashion, is the panel report in EC – Asbestos. The panel, it will be recalled, as much as conceded that the French ban on asbestos in the 1990s could not have been foreseen in 1947 and 1962 when some of the relevant concessions were negotiated. The move that it made in response was to suggest that there could be no "reasonable" expectations of continued market access based on a situation that existed so many years ago – negotiators should recognize that things are likely to change over such a lengthy time horizon.[1]

This aspect of the reasoning in Asbestos seems at odds with the logic of the non-violation mechanism. Of course the negotiators in 1947, 1962 and at every other point in time recognized that unforeseen contingencies might upset the balance of concessions. It is precisely because things do change over time, in ways that are difficult to predict and contract over in advance, that the GATT included the non-violation principle. To refuse to apply it on the grounds that reasonable negotiators would not expect things to stay the same over time seems quite peculiar.

Indeed, the logic in Asbestos could be applied to any context where the negotiations are a number of years in the past. EC – Oilseeds concerned tariff concessions that were more than a quarter century old. They were impaired by the introduction of new European subsidies. Why not invoke the line of reasoning in Asbestos to deny the claim, on the grounds that no reasonable

[1] Contrast Canada – Administration of the Foreign Investment Review Act, BISD 30S/140, adopted February 7, 1984. In that proceeding, a challenge to trade-related aspects of Canada's investment review regime was carefully limited to matters covered by GATT, impicitly suggesting that GATT created no reasonable expectations of any limitation on the sovereign right of members to regulate investment.

negotiator could ever expect Europeans to forego agricultural subsidies over an extended period of years?

The next sections will consider whether there is a more convincing basis for the outcome in cases like Asbestos, and indeed will suggest alternative rationales for it. Suffice it to say here that the case law as it stands offers no sensible account of why the non-violation remedy is so "exceptional."

3. Existing Economic Theory on the Non-Violation Doctrine

Much like the GATT negotiating history from the 1940s, the economic literature on trade agreements suggests an important and potentially expansive role for non-violation complaints. Section 3.1 provides some background on the economic theory of trade agreements, while section 3.2 reviews the existing commentary on non-violation claims.

3.1. Economic Theory of Trade Agreements

The modern economic theory of trade agreements draws heavily on the economic analysis of contracts, which applies generally to the analysis of international law. Trade agreements, like all treaties, are contracts between or among governments. Treaties orchestrate cooperation that enables the signatories to achieve a joint improvement in their perceived "welfare" relative to the absence of cooperation in the presence of international "externalities" (positive or negative effects of policy decisions on the welfare of other nations). Nations tend to ignore externalities when they make choices non-cooperatively because foreigners generally lack efficacy in the domestic political system. The result is economically inefficient outcomes from a global perspective. International law exists primarily to address such externality problems. See generally Sykes (2007).

With respect to trade policy in particular, the economic literature typically posits that governments behave "as if" they are maximizing a "welfare function" that depends on underlying variables that characterize the state of their economies. These welfare functions can depend on a variety of different things, including political considerations that lead the government to give more weight to the well being of particular industries, individuals or constituencies. The one assumption that pervades most of the literature on trade agreements, however, is that governments' perceived welfare increases, all else equal, when their nation's "terms of trade"

improve. The phrase "terms of trade" refers to the price of a nation's exports relative to its imports on world markets. When the goods exported by a nation increase in price, other things being equal, or the goods imported by a nation decline in price, other things being equal, the terms of trade improve. The assumption that governments prefer an improved terms of trade is a natural one – most actors perceive themselves better off when what they are selling becomes more expensive, or what they are buying becomes cheaper.

A standard view among economists is that the international externality associated with non-cooperative trade policy, at least primarily, flows through the terms of trade. For this externality to arise, nations must be "large" enough that their policy choices affect the prices that foreign exporters receive for their goods or services. If the United States increases its tariff on imported wine, for example, the prices that French wine exporters can command for a given quantity of wine sold on world markets may well decline. If so, those exporters are injured by the tariff increase and France's terms of trade decline. When nations make their tariff choices non-cooperatively, they ignore these costs imposed on foreign exporters. As a result, tariffs become "too high" from the standpoint of global efficiency. The same problem arises with other policy instruments that produce negative terms of trade externalities.

When terms of trade externalities lead nations to erect globally inefficient barriers to trade, trade agreements can enhance efficiency (and thus the welfare of participating nations) by orchestrating a mutual reduction in trade barriers. The drafters of trade agreements nevertheless confront many challenges. Among other things, a myriad of policy measures can become barriers to trade. Although the focus of the original GATT negotiations was on tariff cuts, the drafters realized that quantitative restrictions, state-sanctioned import monopolies and state trading enterprises, domestic taxation, domestic regulatory policies, laws against purportedly "unfair" trade practices such as dumping and subsidization, balance of payments measures, and a variety of other policies can also become trade impediments. Much of the GATT text is devoted to rules designed to prevent changes in these other policies from undermining tariff commitments. Likewise, the drafters were aware that the balance of costs and benefits associated with tariff concessions is subject to uncertainty. What appears to be a politically sensible trade concession at one point in time may become a political liability at another due to some unexpected economic shock. Accordingly, the drafters provided for renegotiation of commitments as well as temporary "escape" from them.

Although the GATT anticipated these issues and addressed them to a significant extent, its

drafters were savvy enough to realize that their foresight was not perfect and that they could not negotiate a rule for every possible contingency that they could imagine – in economic parlance, it was not possible to negotiate the "complete contingent contract." Inevitably, various domestic policies would create trade barriers that were not regulated or imperfectly regulated by the GATT text, and changed economic circumstances might render some commitments politically undesirable. The general provisions for renegotiation and escape from commitments afforded a mechanism for members to withdraw any commitments that had become too burdensome, but did little to protect members against a diminution in the value of commitments made by others. The non-violation claims authorized under Article XXIII for "any measure" or "other situation" were plainly introduced in response to this problem.

3.2. "Internalizing the Externality" Through Non-Violation Complaints

Economists have identified a broad range of policy measures that can impede market access, and that are unregulated or weakly regulated by existing GATT/WTO law. As a result, important terms of trade externalities can arise from policies that do not violate the rules. Consider just a few examples.

First, imagine a three-country setting in which country 1 exports some good to countries 2 and 3, and countries 2 and 3 each export the same good to country 1. Each country is large enough for its trade policies to affect the prices received by foreign exporters. Imagine that countries 1 and 2 enter a bilateral agreement to reduce their tariffs on the goods each sends to the other. If the tariff reduction granted by country 1 applies to country 2's exports to country 1 but not country 3's exports to country 1, then country 3 will suffer a worsening of its terms of trade. This outcome can occur in the GATT/WTO system when country 1 and country 2 create a free trade area (such as NAFTA). And even if country 3 receives the same tariff rate as country 2 on its exports to country 1 pursuant to the most-favored-nation (MFN) obligation in GATT Article I (free trade areas are an exception to this obligation), country 3 may still suffer a worsening of its terms of trade because the concession by country 2 to country 1 increases the demand for country 1's exports to country 2, and causes their price to be bid up to the detriment of consumers in country 3. In short, both the advent of preferential trading arrangements that are legal under WTO law, and new bilateral concessions that respect the most-favored nation obligation, can have adverse terms of trade effects on third parties. See Bagwell and Staiger (2002 chapters 5 & 7) and Bagwell and Staiger (2010).

Bagwell and Staiger (2001, 2002 chapter 8) offer a similar analysis of labor and environmental standards. Imagine a trade agreement between two nations pursuant to which they each reduce tariffs on goods exported to each other. The resulting tariffs are then below the levels that each would prefer unilaterally, in exchange for the reduction in tariffs abroad. The constraint on tariff policy gives each nation an incentive to restrict imports in other ways. One possibility is to lower the labor and environmental standards applied to domestic import-competing firms. A lowering of such standards reduces the costs of these firms and enables them to compete more effectively with imports, just as would a higher tariff. A terms of trade externality arises because foreign exporters lower their prices in response. The result is that the importing nation does not bear the full costs of lower labor or environmental standards, and will tend to reduce them inefficiently from a global perspective (a possible "race to the bottom" scenario).

Staiger and Sykes (2011) consider a different regulatory setting in which an importing nation must choose the product standards (health, safety, etc.) applicable to imported goods. Even if such standards cannot discriminate against imported goods relative to domestic goods (the GATT national treatment principle in Article III), the importing nation may set excessively high regulatory standards because of a terms of trade externality. Intuitively, a higher standard increases the costs of imported goods, and foreign exporters lower their price somewhat to avoid losing sales. In that way, some of the regulatory compliance cost is externalized by the shift in the terms of trade.

Finally, Bagwell and Staiger (2002 chapter 9) consider aspects of competition policy. Imagine a government that must decide whether to permit a merger under its competition laws in an import-competing industry. The merger will raise domestic prices, and allow foreign exporters to raise their prices as well, imposing a terms of trade loss on the importing nation. For that reason, the importing nation may inefficiently block such a merger (perhaps it would yield significant efficiencies).

In each of the above scenarios – creating a free trade area, negotiating a bilateral MFN tariff concession, reducing domestic labor or environmental standards, increasing non-discriminatory standards applicable to imports, or blocking a merger – the policies in question are or are likely to be perfectly legal under GATT/WTO rules. Bagwell and Staiger and Staiger and Sykes note, however, that the non-violation doctrine (or a close analogue) affords a possible solution to the externality problem. In principle, a non-violation doctrine might be implemented to require that when nations make policy decisions that affect the terms of trade for others, they

must make commensurate adjustments in their trade policies to restore the terms of trade to its original level, thereby eliminating the "nullification or impairment." In that way, the nation making the policy choice will "internalize the externality" from its decisions. Unsurprisingly, policy choices subject to such a rule will tend to be efficient in relation to the policy making government's own welfare metric. The non-violation doctrine thus works much like a Pigouvian tax in neoclassical economics – a trade law equivalent to the "polluter pays principle."

It bears emphasis that the list of policy scenarios above is hardly exhaustive. It simply reflects some of the scenarios that have received formal treatment in the literature.[2] Other important scenarios involving WTO-legal measures, such as fiscal and monetary policies, exchange rate policies, and so on can easily be imagined in which unilateral policy actions have potentially important terms of trade externalities, and may accordingly be distorted from the global efficiency perspective. Here, too, the non-violation doctrine might in theory play an important role.

Recalling the discussion in section 2, however, the actual role of the non-violation doctrine appears far more modest, and the few cases touching on the types of scenarios above have not been terribly successful from the complainant's perspective. EEC – Citrus raised a non-violation claim in relation to tariff preferences that were ostensibly legal under GATT Article XXIV (pertaining to customs unions and free trade areas), but the panel decision was not adopted by the GATT membership. EC – Asbestos raised a non-violation claim regarding regulations applicable to imported products, but was unsuccessful. Japan – Film challenged, among other things, aspects of Japanese competition policy, but was also unsuccessful. Beyond these few failed or un-adopted challenges, untold numbers of WTO-legal "measures" with potentially important terms of trade effects have not been addressed.

4. A Theory of the Limited Role of Non-violation Claims

A companion paper (Staiger and Sykes 2013) develops a formal model of trade disputes that allows for the possibility of both violation and non-violation claims. It uses the model to explore the conditions under which non-violation claims will be observed, their frequency relative to violation claims, and their success rate. Its goal is to explain the stylized facts noted earlier – the

[2] As we have noted, the introduction of a domestic subsidy is also an important instance where non-violation claims have been used in practice; see Bagwell and Staiger (2006) for a formal analysis of non-violation claims in this context.

selective rate of non-violation filings in the GATT/WTO system, the low rate of adjudication on filed non-violation claims, and the relatively low success rate of those few non-violation claims that are adjudicated – and to provide a platform for assessing the potential value of the non-violation clause under conditions that would lead to these kinds of observed performance measures. Here we sketch the model setup and results in that paper.

4.1. Non-violation Claims "On and Off the Equilibrium Path"

Although the non-violation *doctrine* plays a potentially important role in the economic contributions noted in section 3, these contributions do not necessarily imply that formal non-violation *claims* will play a large role in practice. The difference relates to what economists term behavior "on the equilibrium path" versus behavior "off the equilibrium path." Behavior on the equilibrium path refers to behavior that will be observed in practice. Behavior off the equilibrium path refers to behavior that will not be observed in practice, because the conditions that trigger it do not arise. To give a simple example, imagine that the penalty imposed on individuals who exceed the highway speed limit is boiling in oil. Because of this severe penalty, no one ever exceeds the speed limit, and hence no one is ever boiled in oil. Put differently, "equilibrium" behavior entails no incidents of speeding, and thus no punishment for speeding. Boiling in oil is "off the equilibrium path."

The fact that behavior is off the equilibrium path, however, does not mean that it is unimportant, as the above example makes clear. If the penalty for speeding deters it altogether, then that penalty plays a key role in shaping behavior even though it is never observed. The theoretical treatments of the non-violation doctrine described above are not necessarily inconsistent with the lack of observed non-violation claims in GATT/WTO practice precisely because the role of the non-violation doctrine in these contributions can typically be interpreted to be off the equilibrium path (see, for example, Bagwell and Staiger, 2006, note 6). In particular, these prior contributions discussing the non-violation doctrine arise in theoretical models that do not really incorporate a formal role for observed "disputes" of any kind, whether involving violations or non-violations.

That said, with its weak observed performance measures, we are hesitant to assert that the non-violation doctrine plays an important (off-equilibrium) role in the WTO on the basis of this prior theoretical work. To continue with the highway analogy, if speeding tickets were rarely issued, and when issued were virtually always reduced to a warning upon appeal, it seems

plausible that any deterrent effect of the "boiling in oil" sanction might be minimal despite its nominal presence on the books. Only a richer account of the legal system can illuminate the true importance of the sanction.

Similarly, if economic theory is to be useful in explaining the relative importance of non-violation claims and their success rate, we suggest that a richer account of the WTO dispute process is needed. In particular, theory must suggest a role for a dispute resolution system and predict that disputes within this system are actually observed in practice (in "equilibrium"). One such theory is developed by Maggi and Staiger (2011), but that model does not consider the possibility of non-violation claims. In our companion paper we adopt and extend the model of Maggi and Staiger to incorporate the possibility of non-violation claims, and identify a set of conditions under which the model predicts the above-noted features of observed violation and non-violation claims in the GATT/WTO. We then use the extended model with these conditions in place to consider the nature and potential importance of the role that non-violation claims might play in the GATT/WTO.

4.2. The Modeling Framework

Maggi and Staiger (2011) develop a model of dispute settlement in an international trade agreement with one exporting country and one importing country. There is only one good involved. The importing government has a binary trade policy choice, termed Free Trade or Protection. Protection always harms the exporting country, but whether it benefits the importing country by more than it harms the exporting country – and hence whether Protection or Free Trade is the optimal policy choice from the joint perspective of the parties to the trade agreement – depends on the underlying "state of the world," a configuration of background conditions that is unknown at the time the trade agreement is negotiated.[3] Such conditions might include the quantity of imports that arise under each possible import policy, whether or not a domestic industry suffers severe economic distress, and so on. Ex ante, before uncertainty over the state of the world is resolved, the governments of the importing and exporting country can write an incomplete contract and can also set up a dispute settlement body (DSB) and define its mandate. Ex post, once uncertainty is resolved, the importing government makes its trade policy choice and the exporting government decides whether to initiate a dispute, which

[3]In this model, Free Trade is not necessarily jointly optimal, because for instance government preferences may take account of distributional issues as well as traditional efficiency issues.

if initiated is resolved by the DSB according to its mandate.

Our companion paper extends this basic model. We allow the importing government, in addition to its trade policy choice, to make a domestic regulatory choice, which we term Free Trade or Regulation. And we allow that in some states of the world Regulation is jointly optimal, in other states of the world Protection is jointly optimal and in the remaining states of the world Free Trade (laissez faire) is jointly optimal. The exporting country is always harmed by the trade effects of Protection or Regulation imposed by the importing country. In those states of the world where Protection is jointly optimal (e.g., the imported product is causing substantial injury to domestic workers), the importing country does best with Protection, but it could also benefit from Regulation as a "second-best"/inferior form of intervention relative to Free Trade and may choose Regulation rather than Protection for reasons relating to the dispute settlement system that we detail below. In states of the world where Free Trade is jointly optimal, the importing country benefits from intervention solely as a result of "beggar-thy-neighbor" terms-of-trade consequences and so in these states the importing country does best with Protection, but it could also benefit from Regulation as a second-best means of terms-of-trade manipulation and may make this choice, again for reasons relating to the dispute settlement system that we detail below. Finally, in states of the world where Regulation is jointly optimal (e.g., the imported product poses a health risk) the importing country does best with Regulation; it might also benefit from Protection as a second-best form of intervention in these states, but as we detail below it will never have reason to make such a policy substitution and so we can ignore this possibility.

A first-best trade agreement would require the importing nation to make the efficient choice over each policy instrument given the realized state of the world. Unfortunately, however, the first-best is not achievable because of limitations on the ability of the parties to write an ideal contract. We assume that the realized state of the world (e.g., whether or not the imported product contains asbestos, whether or not imports have surged) is observed both by the governments and by the DSB, but the associated welfare levels of each government (e.g. the degree of health risk associated with an asbestos-containing product, the degree of injury to the domestic industry associated with an import surge) are observed only by the governments and not by the DSB (so that the contract cannot simply provide that the DSB orders the efficient policy choice given the realized state). Still following Maggi and Staiger (2011), we further assume that it is costless to describe the trade policy choice (Free Trade or Protection) in a

contract, but prohibitively costly to describe all the relevant states of the world that may arise, and thus impossible to write the "complete contingent contract" that specifies the optimal trade policy choice in any realized state. Instead, the trade agreement takes the form of what Maggi and Staiger term a "vague" contract that is composed of statements that make Protection permissible under some fairly general conditions, such as "Protection is allowed if and only if increased quantities of imports cause or threaten serious injury to a domestic industry." Vague contracts are costless to write, but their meaning is ambiguous in some states of the world and introduces the possibility of error into dispute rulings. Maggi and Staiger discuss the conditions under which it is optimal for the parties to write such a vague contract, and in effect we simply assume that those conditions will be met here.

Our assumptions about the contracting possibilities over domestic regulation differ from those concerning trade policy. In particular, we assume that writing an ex-ante contract to describe the option of Regulation would be prohibitively expensive, reflecting the notion that regulation can encompass any of a myriad of policies that might be implemented ex post. Hence, even a vague contract cannot be written to cover the choice of regulatory policy.

In addition to writing the ex-ante contract covering trade policy, governments can introduce a court – the DSB – and can give the DSB a mandate to follow if it is invoked to settle a dispute ex post. We consider two possible roles for the DSB - it can address a *violation complaint*, and/or the DSB can address a *non-violation complaint*.

A violation complaint occurs when Home selects Protection as its trade policy and the exporting country claims that the general conditions allowing for protection as stated in the contract do not exist. We focus on states of the world in which ambiguity arises in the contract as to whether the conditions allowing for protection exist, and adopt the following approach to modeling the ambiguity. We assume that the DSB can draw an unbiased but "noisy" signal of the joint welfare associated with Protection, which can be thought of as the outcome of an independent investigation in which the DSB "interprets" the contract. The DSB will rule that Protection is illegal if the noisy signal suggests that Protection is inefficient relative to Free Trade, and will make the opposite ruling if the signal suggests that Protection is efficient relative to Free Trade. We assume in the case of a violation complaint that the importing country always complies with the ruling.

In contrast to a violation complaint, a non-violation complaint does not involve a claim that a contractual obligation has been violated, and we assume (in line with GATT/WTO

practice) that a non-violation claim can be brought either when the importing government chooses Protection as its trade policy, or when it selects Regulation as its regulatory policy. If the DSB is asked to rule on a non-violation complaint, it again observes a noisy signal of the joint payoffs from the policy at issue, and based on that signal will rule on the legality of Protection or Regulation. We assume (again in line with GATT/WTO practice) that the importing country is under no obligation to change its policy in the event of an adverse non-violation ruling.[4] Instead, it has the option of either implementing the DSB policy determination or paying *damages* to the exporting country.[5] We assume that the damages *paid* by the importing country are equal to the value of the harm suffered by the exporting country. We also assume that the exporting country *receives* less than the importing country pays, reflecting the dead-weight loss associated with tariff retaliation, the typical form of retaliation in GATT/WTO disputes.[6] And if both a violation and a non-violation claim are brought against the importing government's choice of Protection, we follow GATT/WTO practice and assume that the violation claim is first ruled upon, and the non-violation claim is subsequently ruled upon only if the violation claim was unsuccessful. Finally, we assume that disputes are costly. If the exporting country (complainant) invokes the DSB, each country must pay a litigation cost. A non-violation claim in addition to a violation claim requires additional litigation cost.

4.3. Equilibrium Behavior

The analysis of this model yields a number of interesting results, which we will simply state here. The reader is referred to our companion paper for proofs.

Our assumptions about the contracting environment make it optimal under certain conditions for the parties to create a dispute settlement mechanism and to utilize it. The fact that the DSB does not always reach the right answer due to its noisy signal about the efficiency of the importing country's policies leaves room for disputes to arise in equilibrium. Even though the two governments are assumed to know whether the policy at issue is efficient, the fact that the DSB may err can induce the exporting country to file a complaint against an efficient policy,

[4] See DSU Art. 26.1.b.

[5] There is an important question as to the practical distinction between violation and non-violation complaints in the GATT/WTO, in light of the fact that the same reciprocity-of-trade-effects rule generally guides the permissible retaliation for continued application of the intervention at issue in either case. On this question there is a debate among legal scholars (see Jackson, 1997, and Schwartz and Sykes, 2002), but the model abstracts from this debate in order to focus on other issues.

[6] Maggi and Staiger (2012) provide some discussion of the methods of compensation typically available in GATT/WTO disputes and the deadweight loss associated with these methods.

and can induce the importing country to try and get away with an inefficient policy. When either of these behaviors arises, a dispute erupts in the model.

The decision by the exporting country to file a case turns on a simple calculus that compares the expected benefits of the case to its costs. Relevant parameters in this regard include the probability of success, the benefits of inducing compliance in a violation case, the amount of damages received in a non-violation case, and the costs of litigation. A violation case is more attractive to the complainant because it results in compliance if it is successful, while the non-violation case may only result in "damages" that are less than fully compensatory due to their deadweight costs (although in some cases the importing country may prefer to withdraw its policy rather than pay damages, namely, cases in which the policy is inefficient). Non-violation claims nonetheless play a role in equilibrium for two reasons. First, with respect to trade policy (which is contractible under the vague contract), a non-violation claim can give the complainant a second chance at success, which is valuable in the event that the violation claim fails. Second, the non-violation claim is the only viable claim against Regulation, which by assumption is not contractible even under a vague standard.

In this environment, disputes will be more frequent when the DSB is *less* accurate. As we have indicated, when a dispute arises it is because one of the parties is acting opportunistically to exploit the incompleteness of the vague contract and the inaccuracy of the DSB rulings. An immediate implication of this observation is that the clearest efficiency-enhancing role of the DSB in this setting occurs off-equilibrium – when its presence deters inefficient conduct. Likewise, in any state of the world in which a dispute arises, behavior is inefficient. One of the parties is behaving opportunistically and inducing costly litigation that would not arise in a first-best world.

We stated at the outset that one of our goals with this model was to offer an explanation for certain stylized facts – non-violation claims are infrequent, are often not adjudicated, and those that are adjudicated are not often successful. Under plausible parameter values pertaining to the magnitude of litigation costs, the accuracy of DSB rulings, the deadweight costs of "damages" in non-violation cases, and the relative benefits to the importing country of Protection versus Regulation as a means of terms-of-trade manipulation, the model predicts that (a) the probability of a successful ruling in a violation case will tend toward the probability that the DSB will rule accurately, and that (b) the probability of a favorable ruling in a non-violation case will tend toward the probability that the DSB will rule inaccurately. Hence, if the DSB

is reasonably (but not perfectly) accurate, violation claims will have a high success rate and non-violation claims will infrequently produce a favorable ruling. Part of the reason is that non-violation claims will not be filed very often unless the policy being challenged is inefficient, and in the bulk of those cases a violation complaint will also be filed and succeed, making a non-violation ruling unnecessary. Likewise, because the most meritorious non-violation claims are often not adjudicated for this reason, the success rate for non-violation claims that do receive a ruling will be lower than for violation claims.

As we noted, a second goal with this model was to use it to assess the potential value of non-violation claims. Here we find that, despite the limited role of non-violation claims "on the equilibrium path," and their relatively lower success rate, they can still play an important role in the system. In the model, the absence of a non-violation remedy would imply that the importing nation could choose Regulation with impunity, even when it is inefficient. And as we indicated at the outset, without the ability to bring a non-violation claim the importing nation may indeed choose to intervene with regulation when it would not be jointly optimal to do so (either when Protection is jointly optimal, or when Free Trade is jointly optimal) in order to avoid a dispute over Protection. The non-violation claim provides a check on such behavior. It also provides a "backup" option for challenging Protection, that is valuable when Protection is inefficient but a violation claim may fail. In both cases, non-violation claims play a role both off equilibrium (in deterring the inefficient policy choice) and, less frequently, in equilibrium (resulting in an observed filing). The downside of the non-violation mechanism is that it may be used opportunistically by the exporting nation to challenge an efficient policy. But once again, under plausible parameter values that comport with those necessary to generate the pattern of results we observe in the GATT/WTO system, this (observed, on-equilibrium) use of the non-violation complaint will be rare, and the net effect of the non-violation mechanism will therefore be favorable.

5. Going Beyond the Model: Further Thoughts on the Limited Role of Non-Violation Claims

Our model, coupled with plausible assumptions about pertinent parameter values, offers a possible explanation for the selective filing rate for non-violation claims in the WTO system and the low success rate for such claims. It also suggests that the non-violation doctrine can play an important role nevertheless. But a number of issues remain unexplored. For example,

is there any relationship between non-violation claims and the distinction in international law between good faith and bad faith performance of treaty obligations? Why has the role of non-violation claims diminished over time, despite the importance envisioned for them by the GATT drafters? Why have successful non-violation claims in the system been limited to complaints about "commercial" measures? To what extent are the market access expectations protected by the non-violation doctrine time-limited? This section offers some informal thoughts on these and related questions.

5.1. "Bad Faith" and Private Information

Consider a trade negotiation in which a trade negotiator has private information about some definite or likely future policy change by her government. The policy change will significantly diminish the sales opportunities for foreign exporters of a good or service that is the subject of the trade negotiation. The negotiator does not reveal this information, and proceeds to negotiate a concession on imports of the good or service. In a more extreme version of this scenario, one might even imagine that the negotiator has affirmatively misled her counterparts about the likelihood of the policy change. In a less extreme version, one might imagine that the negotiator's superiors have deliberately or carelessly kept the negotiator in the dark about the likely policy change. In any case, after the conclusion of negotiations, the policy change is enacted, and foreign exporters are disappointed that the concession they received is considerably less valuable to them than expected.

In this class of scenarios, the negotiator (or her superiors) with private information has engaged in deceit or opportunistic non-disclosure. Any reciprocal concession by counterparties has effectively been obtained by inducing false expectations about what the counterparties would receive in return, and the negotiating country may fairly be said to have acted in bad faith.

If we assume further that the policy change at issue is legal under WTO law, this scenario suggests the clearest case for a non-violation claim. True, negotiating counterparties could always ask for affirmative representations about all potential policy changes and memorialize them as formal commitments in schedules of concessions, but the process of identifying, discussing and memorializing such matters would be cumbersome, costly, and no doubt incomplete. The generic non-violation claim is likely the cheaper way of dealing with the problem, provided of course that evidence of bad faith can be produced to prove the claim. Such a showing is

likely to be possible in many cases because changes in government policy are often preceded by bureaucratic activity that generates a paper trail back to the source, particularly in the democracies.

There is no reason why non-violation claims relating to bad faith should be limited with regard to the nature of the policy change at issue. Even if the policy involves a new regulation with the bona fide purpose of preserving human life, for example, the importance of the pending regulation is no excuse for concealing it during the negotiation process.

Although bad faith scenarios present the most compelling cases for non-violation claims, none of the non-violation disputes in the history of the GATT/WTO system reflect clear evidence of bad faith. Perhaps the closest is Korea – Procurement, in which Korean negotiators may have (perhaps accidentally) misled U.S. counterparts about the agency that would be responsible for the airport project in question, but even then the legislation specifying the responsible entity was on the books and available for the United States to examine before the close of negotiations over the GPA.

The paucity of cases raising serious suspicions of bad faith, however, does not mean that the non-violation doctrine plays no role in policing such behavior, as our formal analysis (albeit not with private information) has illustrated. Because cases of demonstrable bad faith present such a compelling case for a successful claim, negotiators and their superiors may simply be dissuaded from bad faith behavior. The doctrine may thus have an important effect on the conduct of negotiations even though actual incidents of bad faith are off the equilibrium path.

5.2. Contractibility as a Function of Time

In our model, we assumed that "tariffs" were contractible (at least in a "vague" way) and "regulation" was not. This stylized assumption captures an important feature of treatymaking in practice. Some issues are readily reduced to simple rules (the tariff on widgets shall be no more than 10% ad valorem), some can be made subject to vague standards (concessions may be withdrawn if domestic producers of like or directly competitive products experience serious injury as a result of increased quantities of imports), and some are too difficult even to subject to vague standards (macroeconomic policies perhaps?). In each setting, the details of treatymaking will be a function of the complexity of the issue and the attendant costs and benefits of fashioning rules or standards to address it.

If the benefits of creating treaty obligations increase and the costs decline with time, we

might expect parties to respond by creating additional rules and standards. With reference to the WTO in particular, there are two reasons to think that the costs of writing efficiency-enhancing treaty commitments have declined in relation to the benefits. First, the WTO has greatly expanded its membership over time, and the volume of trade subject to GATT rules has increased dramatically in real terms. As a trade agreement grows in size and the volume of trade grows as well, the number of instances in which any efficiency-enhancing principle applies will likely increase. But the cost of creating a treaty commitment may be roughly fixed (in real terms). If so, the benefits of a given efficiency-enhancing commitment will tend to increase in relation to its cost.

Second, the WTO membership has accumulated decades of experience with the problems posed by changing circumstances. This experience likely allows the membership to make better judgments about which changes in conduct are efficient and which changes are not. Likewise, where the efficiency of some change in conduct depends on background conditions, the membership has gained experience in identifying the background conditions that make a particular change efficient or inefficient. The accumulation of such experience directly reduces the costs of writing efficiency-enhancing rules and standards, and tends to increase their benefits as well by making them more accurate.

Accordingly, it is no surprise that the history of the WTO shows a steady progression of legal detail. The 1947 GATT was subject to significant amendment during the 1950s review sessions. The 1960s brought the early supplemental codes (such as the first Antidumping Code), which led to a number of more extensive codes during the 1970s Tokyo Round. Finally, the vastly elaborated treaty text and revised codes of the WTO grew out of the Uruguay Round negotiations in the late 80s and 90s.

This process has important implications for the utility of the non-violation doctrine. At the outset of GATT, when the costs of crafting rules and standards were higher in relation to their benefits, considerably fewer rules were devised. Many contingencies were not specifically addressed. The potential value of a non-violation doctrine as a fallback, to assist in policing potentially inefficient behavior that was not covered by explicit rules, was likely at its peak. Over time, however, as more rules and standards were crafted to delineate violations (i.e., inefficient behavior), the number of situations left for the non-violation doctrine to police has plausibly declined. One might view the rule in the WTO Subsidies Code that makes a new and unexpected subsidy "actionable" if it impairs a negotiated concession, for example, as a

nice illustration of how newly elaborated "violations" can take an issue out of the non-violation realm.

Likewise, as the number of rules and standards has expanded, the membership has at least implicitly made a collective decision that certain changes in conduct are efficient even if they may have adverse terms of trade externalities. This inference is particularly plausible where a set of issues has been considered at length in the negotiating process, and detailed rules have emerged to govern them.

Such observations offer some support for a species of argument advanced by the European Union, and rejected by the Appellate Body, in EC – Asbestos. Europe argued that non-violation complaints should not be allowed if the measure at issue was legal under a specific GATT provision. The rules governing regulatory measures for health-related and other purposes were addressed somewhat in the original GATT via the Article III national treatment obligation and the Article XX exceptions. These issues were the subjects of further negotiations during the Tokyo Round, which produced the Tokyo Round Code on technical barriers. Subsequent extensive negotiations during the Uruguay Round led to the WTO Agreements on Technical Barriers to Trade (TBT) and Sanitary and Phytosanitary Measures (SPMs) (only the former was applicable in EC – Asbestos). Especially if the measure at issue in Asbestos had passed muster under the TBT Agreement – an issue that was not adjudicated for reasons that need not detain us – one might argue that the non-violation complaint ought be foreclosed, or at least that the measure at issue should benefit from a strong presumption of efficiency.

The more general point is that as increasingly detailed rules have evolved since the founding of GATT in 1947, one must ask whether conduct that is permitted under the rules is presumptively efficient and thus an inappropriate subject for a non-violation complaint. On this point, however, our model does suggest some caution: as we have observed, a potentially important off-equilibrium impact of the non-violation clause is that non-violation claims may provide a back-up complaint that can be used to beef up a claim of alleged contractual violations, and thereby make such violations less attractive; and clearly such an impact would be destroyed if the scope of non-violation claims were limited to non-contracted policies. Nevertheless, in a richer model it might be possible to identify cases in which the failure of the violation complaint ought also be fatal for any potential non-violation complaint. For the reasons noted here, we suspect that this could be the case for at least certain types of behaviors that have been scrutinized closely in detailed negotiations.

5.3. Non-Violation Claims as a Liability Rule Mechanism: Some Practical Issues

Our model assumed that a successful violation claim would be followed by compliance, while a successful non-violation claim would be followed by a "damages" payment. The assumption that respondents always comply with a violation ruling hints at an underlying "property" rule, whereby violations are never profitable and deviation from commitments will only be possible after securing permission from affected counterparties. The practice of paying "damages" after a non-violation claim, by contrast, suggests an underlying "liability" rule, whereby the option to continue the challenged behavior arises as long as the respondent pays the stipulated price.

These features of the model touch indirectly on a lively debate among WTO scholars. Numerous economic commentators have argued that the WTO system exhibits certain characteristics of a liability rule system. Under GATT Article XXVIII, for example, members that wish to withdraw tariff commitments must initially negotiate for the right to do so, but if they are unable to negotiate a satisfactory outcome, they may proceed to withdraw the commitment, and adversely affected members may thereafter withdraw "substantially equivalent concessions" of their own. To the degree that "substantial equivalence" implies a price for withdrawal that approximates the harm done to others (a debatable proposition in practice to be sure), Article XXVIII has a liability rule quality about it. The WTO Dispute Settlement Understanding (DSU) introduced this principle into disputes under GATT Article XXIII. In both violation and non-violation cases under the DSU, complainants are entitled to withdraw "substantially equivalent concessions" when the target of a complaint does not either remove the objectionable measures or offer adequate compensation for them. A neutral arbitrator is charged with assessing "substantially equivalence" should the disputants disagree. Once again, a liability rule approach appears to be operative. See, e.g., Schwartz & Sykes (2002)

This suggestion that the WTO dispute settlement system creates a liability rule mechanism fits nicely with the economic commentary, discussed earlier, on the function of non-violation claims. A liability rule, with a properly computed compensatory "price" for changes in conduct that affect others adversely, can in principle induce trading nations to internalize the terms of trade externalities from their policies.[7] If the price for deviation is set efficiently, and a violator

[7] It is worth emphasizing that such a mechanism can have attractive features even when the aggrieved parties don't actually receive full compensation for the harm suffered. It is enough that when a party alters its conduct in a manner that imposes harm, that party incurs a cost that approximates the harm. Thus, even if the "price" to be paid by a losing disputant takes the form of trade retaliation by other nations, and even if retaliation does not restore the welfare of those parties to its original level (although in principle it might), the system can

prefers to pay the price rather than comply, then deviation is presumptively efficient.

An important difficulty with a liability rule mechanism, however, is that its efficiency depends critically on the ability of adjudicators to set a proper price ex post for conduct that causes harm. In this regard, our model assumed that the "damages" were at least somewhat undercompensatory (reflecting in part the deadweight loss associated with tariff retaliation as the form of damage payment), and the possibility of a significant gap between harm suffered and damages payable plays an important role in explaining the pattern of disputes.

Moving beyond the model, the task of operationalizing the "substantial equivalence" principle in WTO practice is a daunting one. The welfare loss to complaining nations and the loss imposed by trade retaliation on its target is not simply a monetary loss as in a tort or contract suit, but a political loss that is not directly verifiable and, by hypothesis for purposes of the present discussion, is poorly correlated with other verifiable information. Even if adjudicators view their task as one of implementing an efficient liability rule, therefore, it is difficult to see how they will avoid large errors. Furthermore, as reflected in our model and mentioned just above, the "damages" payable in the WTO carry a deadweight loss because they take the form of trade sanctions, which distort trading volumes inefficiently. The costs of litigating over the proper amount of retaliation before WTO arbitrators add further to this problem. As a result, the calibration of the efficient penalty becomes even more complex and the resulting incentive structure can be no better than "second-best."

For such reasons, various commentators express skepticism about the utility of a liability rule mechanism in an organization like the WTO, and argue that a property rule approach may be better despite its own imperfections. See generally the essays in Bown & Pauwelyn (2010). The skeptics see little hope that the "substantial equivalence" principle can serve to optimize ex post incentives to deviate from WTO rules, and instead favor the negotiation of detailed rules ex ante backed by a property rule sanction designed to force renegotiation by any party that wishes to deviate.

Even if the skeptics are right, of course, it remains puzzling as a positive matter as to why WTO negotiators settled on the "substantial equivalence" rule in preference to some stiffer level of penalty that would better implement a property rule approach. A possible answer is that occasional situations will inevitably arise in which negotiated rules are inefficient, and deviation

still help to encourage efficient adaptation to changing circumstances. On the other hand, as we have earlier observed and next discuss further, our model also assumes inefficiencies associated with the damage payments, and this does complicate and qualify the standard efficient breach interpretation.

may be quite important. A property rule may make it impossible for nations to deviate under those circumstances and threaten to cause the trading system to unravel altogether. In short, if the consequences of forbidding efficient deviation are more serious than the consequences of allowing inefficient deviation, a liability approach in *violation* cases may still be preferable to a property rule.

In non-violation cases, by contrast, the option exists to have no remedy at all, leaving the externality problems caused by conduct that is legal to future negotiations. This option may become attractive for the reasons enumerated above – when the "price" to be paid following a successful non-violation complaint does not capture the harm done to others with much accuracy, and when the other costs of the system including the economic costs of trade sanctions and the costs of litigation are substantial, one must then consider the possibility that the game is simply not worth the candle.

In this regard, it is noteworthy that the challenges of identifying "equivalent" concessions have likely risen considerably since the founding of GATT. In 1947, initial tariff concessions were exchanged, and it may well have been possible in many instances to unpack what was exchanged for what. With the passage of decades since many concessions were negotiated, and after numerous additional negotiating rounds, the capacity to identify the "reciprocal" concessions in any given case and the anticipated effect of particular concessions is for the most part gone. In actual WTO disputes, the arbitrators do not even try to identify "substantial equivalence" based on the negotiating history, but instead tend to permit retaliatory measures that impose a loss of trade volume on the target roughly equal to the loss of trade volume due to the measure that was the basis for the complaint. The linkage between a sanction based on lost trade volume, and the level of sanction that would (in theory) encourage efficient adaptation to changing circumstances, is tenuous. See Grossman & Sykes (2010).

A further consideration, noted in our presentation of the model and in the last section, is that WTO members may have negotiated rules and standards to govern the most important forms of inefficient behavior, leaving less and less of a role for the non-violation doctrine to police serious inefficiencies. If the most important matters are already covered by the commitments that underlie violation claims, and the liability rule mechanism that governs non-violation claims is quite error prone and costly for the reasons noted, we have further justification for the receding significance of non-violation claims. Nevertheless, here our model again suggests some caution, as it indicates that the off-equilibrium value of maintaining the possibility of bringing

non-violation claims in the WTO could be substantial even as the importance of observed non-violation cases recedes.

5.4. Regulatory Chill?

In standard economic models of trade agreements (e.g., Bagwell & Staiger, 1999, 2001), national governments are assumed to represent the interests of their own constituents faithfully. The role of a trade agreement is simply to address the international (terms of trade) externality, and once that externality has been addressed properly, subsequent policy decisions will be globally efficient in light of the objectives of the member governments. As discussed above, the posited role of the non-violation doctrine in these standard models is to induce the internalization of terms of trade externalities in the face of contractual incompleteness.

Popular criticism of trade agreements, in contrast, often focuses on their purported threat to aspects of "sovereignty." Opponents of the WTO and various preferential trading arrangements, for example, have often contended that such agreements discourage desirable domestic regulatory measures to protect human health, the environment and other important matters.

It is possible to square such concerns with the standard models of trade agreements (see e.g., Bagwell and Staiger, 2002, Ch 8), but it is also true that those models assume away a potentially important class of issues that may also be relevant to an understanding of the issue. In this regard, the broader public choice literature raises a number of concerns about the tendency of governments to favor some domestic constituencies over others, perhaps to the detriment of the nation as a whole. If imperfections in the political process lead governments to sacrifice the interests of some group for another, even when the interests of the former group are in some "objective" sense more important, measures to correct international externalities may introduce other distortions – an application of the economic theory of the "second-best."

Thus, perhaps the "sovereignty" objections to trade agreements can be recast slightly. Suppose, for example, that national governments under-regulate product health and safety because such regulations tend to impair the interests of large and well-organized producer groups, and benefit less well-organized consumers. Then, to the degree that trade agreements further discourage product health and safety regulation (due to such things as national treatment requirements, requirements in the technical barriers agreements, and perhaps non-violation complaints), such regulation becomes even more sub-optimal. WTO rules in that sense might create "regulatory chill." Conceivably, these distortions might impose greater costs than the

benefits gained from addressing terms of trade externalities.

Of course, one cannot make much of such domestic political imperfections in fashioning a positive theory of trade agreements, because governments that undervalue the interests of certain constituencies in making regulatory policy will also presumably undervalue them in international negotiations. But these concerns may nevertheless afford some basis for normative critique, and that critique may come to play a political role in how the trading system evolves over time if its proponents become better organized.

Consider, once again, a case such as EC – Asbestos. Assume that the French prohibition on asbestos-containing products comports with all pertinent obligations in the treaty text (not fully adjudicated in the actual case, as noted earlier), and imagine that the panel or the Appellate Body had nevertheless found that the French measure results in non-violation nullification or impairment, obliging France to provide compensation or suffer retaliation. It is easy to imagine that such a decision would have become a cause célèbre for WTO critics, and that their critique would largely track the above narrative – product health and safety standards are often too low to begin with, and WTO rules make them worse.

At a minimum, these observations offer a possible reason why WTO adjudicators – who are not entirely insensitive to the politics of the system – may have been reluctant through the years to invoke the non-violation doctrine when the measure at issue relates to politically sensitive issues such as human health, the environment, and conservation of natural resources. Likewise, they may suggest a reason why member governments have generally been reluctant even to assert non-violation claims relating to such matters. Further, depending on one's view of the underlying normative issues, the reluctance of adjudicators and members to invoke the non-violation doctrine in such cases may be desirable as a normative matter.

By contrast, recall the cases where non-violation claims have succeeded – new farm subsidies, discriminatory farm subsidies, and tariff discrimination among competitive products. None of these policies seem to raise obvious concerns about regulatory chill. Instead, such "commercial" policies simply have the effect if not the intent of providing some enterprises with a competitive advantage and do not seriously implicate broader policy values that might seem underserved by existing regulation.

Thus, notwithstanding the open-ended treaty text, the historical limitation of non-violation claims to "commercial measures" may be explicable as a political matter, and may have plausible normative justification as well as long as there has been no bad faith in the negotiation

process. Likewise, although both the panel and the Appellate Body in EC – Asbestos rejected the proposition that the non-violation doctrine should be cabined to "commercial measures," a plausible case can be made for a presumption against its applicability outside of the commercial area (absent evidence of bad faith). Here again, though, our model gives us some cause for caution, or at least for an alternative interpretation: as our model results make clear, the most likely observed (on equilibrium) use of the non-violation claim will be in circumstances where the "regulatory" policy is most substitutable for "protection" as a means of terms-of-trade manipulation, and one could readily imagine that this relatively high degree of policy substitutability will occur precisely for "commercial measures" such as subsidies and discriminatory fiscal instruments.

5.5. An Expiration Date for Market Access Expectations?

If the function of the non-violation doctrine is to induce national governments to internalize the externalities from changes in their conduct, there is no apparent reason to limit it to cases in which changes in conduct impair the market access expectations associated with recent tariff negotiations. In other words, there is no reason why protection for market access expectations should "expire" with the passage of time. Whether a nation contemplates a change in its conduct six months after a negotiation, or sixty years after, it is still valuable to ensure that the international externalities are internalized.

Indeed, if the doctrine really works well to induce the internalization of terms of trade externalities, there is no reason to limit it only to situations where a specific concession has been impaired. Recall, for example, EEC – Citrus, which involved tariff preferences given by the EEC to certain favored exporters in the Mediterranean basin. Even if the tariffs on those products were unbound (no negotiated concessions existed), the introduction of discrimination imposed terms of trade externalities on nations that did not receive the preferences. The resulting pattern of trade may have been globally inefficient, and the non-violation doctrine might in principle discourage such inefficient deviations from the MFN principle.[8]

Our discussion above, however, offers reasons to be pessimistic about the ability of an open-ended non-violation doctrine to police such conduct very well. The capacity of adjudicators to

[8] That said, aside from instances of discriminatory policy such as the one we have described in the text, the standard theory of trade agreements would provide support for limiting the reach of the non-violation doctrine to situations where a specific concession has been impaired, because according to the standard theory in the absence of such a concession there would be no motive for inefficient domestic policy choices.

measure the harm imposed on others and translate it into an appropriate sanction, the direct economic costs associated with trade sanctions, and the substantial costs of litigation all call into question whether extensive use of the non-violation mechanism will yield net benefits.

One might then imagine a more modest role for the doctrine, which focuses more narrowly on its contribution to facilitating the exchange of tariff concessions. Reflecting back on the negotiating history of GATT from the 1940s, the negotiators were not wringing their hands over possible terms of trade externalities at the dawn of the new millennium. Their time horizon was much shorter, and their focus was on the possibility that unanticipated developments in the near term post-war economy might pull the rug out from under the negotiated balance of concessions.

From this perspective, the function of the non-violation doctrine is not to ensure efficient adaptation to changing circumstances over an unlimited time horizon, but to create a comfort level in trade negotiators that will enable them to exchange more concessions during a negotiating round. Because negotiators will discount the future, events many years later will be of little consequence to them and will have little impact on their willingness to exchange concessions. And as noted earlier, the implementation of the non-violation doctrine may be somewhat easier following a recent trade negotiation if the negotiators can identify which concessions have been exchanged for which, and adjust the bargain accordingly should it prove inequitable due to some unanticipated near term shock.

As a consequence, an argument can be made for limiting non-violation protection to changes in conduct that impair a relatively recent bargain. Negotiators then have some confidence that they will not be seen to have "given away the farm" in the event of an adverse shock in the near future. Possible shocks in the distant future will be of little concern to them and the absence of a non-violation mechanism to address them will not discourage trade-liberalizing bargains.

To be sure, it may not always be clear whether a particular issue affects a recent bargain. Even if the tariff on widgets remains fixed during a negotiating round, for example, negotiators might claim that their willingness to accept the bargain on other matters rested importantly on an expectation that market access for widgets would remain the same. But the doctrine might at least be modified to require a claimant to put forward such an argument, rather than allowing expectations from negotiations in the distant past to form the predicate for nullification or impairment.

This analysis suggests an alternative rationale for discounting aged expectations in a case

such as EC – Asbestos. Recall that the Asbestos panel (in the late 1990s) rejected a non-violation claim based on expectations from the 1940s and 1960s, arguing that negotiators could not reasonably expect regulation to remain the same for such a lengthy period of time. One might instead simply say that expectations from such distant negotiating rounds are no longer protected, because such protection is unnecessary to facilitate trade concessions. Any non-violation claim would have had to be predicated upon reasonable expectations engendered by the more recent Uruguay Round, by which time – as the panel rightly noted – the health hazards of asbestos were well known and prohibitions on asbestos products were already proliferating in the global economy.

6. Conclusion

This paper addresses the gulf between the narrow role of the non-violation nullification or impairment doctrine in historical GATT/WTO practice, and the considerably more expansive role for the doctrine seemingly envisioned by the drafters of GATT and suggested by existing economic theory. Our formal analysis offers a possible explanation for two stylized facts: the selective use of non-violation claims relative to violation claims, and the low success rate for non-violation claims as compared to that for violation claims. Our analysis also suggests how the non-violation doctrine may play an important role in policing opportunism both on and off the equilibrium path even if observed claims are few in number and rarely successful in practice. We also provide an informal discussion of considerations outside the model, and here we emphasize that the non-violation doctrine may play an especially important role in policing bad faith during the negotiating process. But that discussion raises doubts about the ability of the non-violation doctrine to play much more of a role due to the shrinking set of issues left out of the negotiated rules and standards for violation cases, and the costs and errors associated with efforts to calibrate "damages" in the WTO system; it offers some possible reasons why non-violation claims should prove particularly difficult when they concern various "non-commercial" and politically sensitive issues; and finally, it suggests a possible rationale for allowing the protection for market access expectations associated with a negotiating round to "expire" with time.

7. References

Bagwell, Kyle and Robert W. Staiger (1999). "An Economic Theory of GATT," *American Economic Review*.89(1), pp. 215-48.

Bagwell, Kyle and Robert W. Staiger (2001). "Domestic Policies, National Sovereignty and International Economic Institutions," *Quarterly Journal of Economics* 116: 519-562.

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

Bagwell, Kyle and Robert W. Staiger (2006). "Will International Rules on Subsidies Disrupt the World Trading System?" *American Economic Review*. June: 877-895.

Bagwell, Kyle and Robert W. Staiger (2010). "Backward Stealing and Forward Manipulation in the GATT/WTO," *Journal of International Economics* 82: 49-62.

Bown, Chad and Joost Pauwelyn (eds.) (2010). **The Law, Economics and Politics of Retaliation in WTO Dispute Settlement** (Cambridge University Press).

Busch, Marc and Eric Reinhardt (2000). "Bargaining in the Shadow of the Law: Early Settlement in GATT/WTO Disputes", *Fordham International Law Journal* 24:1-2 (November-December), 158-172.

Cho, Sungjoon (1998). GATT Non-Violation Issues in the WTO Framework: Are They the Achilles' Heel of the Dispute Settlement Process?, *Harvard International Law Journal* 39: 311-55.

Cottier, Thomas and Krista N. Schefer (1997). "NVNI Complaints in WTO/GATT Dispute Settlement: Past, Present and Future," in Ernst-Ulrich Petersmann (ed.), **International Trade Law and the WTO/GATT Dispute Settlement System** (The Hague: Kluwer).

Gene M. Grossman and Elhanan Helpman (1995). "Trade wars and trade talks," *Journal of Political Economy* 103(4): 675-708.

Gene M. Grossman and Alan O. Sykes (2010). "Optimal Retaliation in the WTO – A Commentary on the Upland Cotton Arbitration," *World Trade Review* 10: 133-64.

Hudec, Robert E. (1990). **The GATT Legal System and World Trade Diplomacy** (Salem, NH: Butterworth).

Hudec, Robert E. (1993). **Enforcing International Trade Law: The Evolution of the Modern GATT Legal System** (Salem, NH: Butterworth).

Jackson, John H. (1969). **World Trade and the Law of GATT** (New York: Bobbs-

Merrill).

Jackson, John H. (1997). "The WTO Dispute Settlement Understanding–Misunderstandings on the Nature of Legal Obligation," *The American Journal of International Law* 91 (1), January, pp. 60-64.

Johnson, Harry (1953). "Optimum Tariffs and Retaliation," *Review of Economic Studies* 21: 142–153.

Maggi, Giovanni and Robert W. Staiger (2011). "The Role of Dispute Settlement Procedures in International Trade Agreements," *Quarterly Journal of Economics* 126, 475–515.

Maggi, Giovanni and Robert W. Staiger (2012). "Trade Disputes and Settlement," mimeo, December.

Mitchell, Andrew D. (2006). "Good Faith in WTO Dispute Settlement," *Melbourne Journal of International Law* 7: 339-73.

Panizzon, Marion (2006). **Good Faith in the Jurisprudence of the WTO** (Oxford: Hart Publishing).

Petersmann, Ernst-Ulrich (1997). **The GATT/WTO Dispute Settlement System** (London: Kluwer).

Schwartz, Warren F. and Alan O. Sykes (2002). "The Economic Structure of Renegotiation and Dispute Settlement in the World Trade Organization," *Journal of Legal Studies* 31: S179-S204.

Staiger, Robert W. and Alan O. Sykes (2011). "International Trade, National Treatment and Domestic Regulation," *Journal of Legal Studies* 40(1): 149-203.

Staiger, Robert W. and Alan O. Sykes (2013). "How Important can the Non-Violation Clause be for the GATT/WTO?," mimeo, June.

Sykes, Alan O. (2007). "International Law," in **Handbook of Law and Economics**, Vol. I (A. Mitchell Polinsky & Steven Shavell eds.) (Elsevier).

Sykes, Alan O. (2010). "Optimal Sanctions in the WTO: The Case for Decoupling," in **The Law, Economics and Politics of Retaliation in WTO Dispute Settlement**, Chad Bown & Joost Pauwelyn eds. (Cambridge Press).

Sykes, Alan O. (2011). "The WTO Dispute Resolution Mechanism: Ensuring Compliance?," in Martin J. Daunton, Amrita Narlikar, and Robert M. Stern (eds.), **The Oxford Handbook on the World Trade Organization** (Oxford University Press, forthcoming).

WTO (1995). **Guide to GATT Law and Practice** (Analytical Index) 6th ed., Vol. II (Lanham, MD: Bernan Press).