

Working Paper No 2010/10| May 2010

Escalation Dynamics in WTO Disputes Over Environment, Health and Safety Issues

Thomas Sattler, Gabriele Ruoff and Thomas Bernauer

Abstract

The quantitative literature on WTO dispute settlement explains dispute outcomes primarily through country characteristics and neglects policy-variables. The environmental policy literature tends to assume that disputes over inter-jurisdictional differences in environment, health and safety (EHS) regulation are particularly prone to escalation – without testing this assumption for the population of WTO disputes. We focus on these gaps by testing the claim that disputes over EHS regulation are more likely to escalate than disputes over other issues. The theoretical underpinning of this hypothesis is that, relative to non-EHS disputes, gradual concessions by the defendant to the plaintiff and compensation to domestic interest groups in the defendant country in exchange for international trade concessions are harder to achieve in EHS cases. We develop and test this hypothesis with data on 506 dyadic WTO trade disputes in 1995-2003, using selection models. The results show that, *ceteris paribus*, EHS disputes are less prone to escalation from the consultation to the Panel / Appellate Body level, but more prone to escalation into compliance disputes once they have reached the Panel / Appellate Body level. This finding suggests that escalation dynamics at different stages of the WTO dispute-settlement process differ across EHS- and non-EHS disputes and should be studied in greater detail in further research.

NCCR TRADE WORKING PAPERS are preliminary documents posted on the NCCR Trade Regulation website (<www.nccr-trade.org>) and widely circulated to stimulate discussion and critical comment. These papers have not been formally edited. Citations should refer to a “NCCR Trade Working Paper”, with appropriate reference made to the author(s).

Escalation Dynamics in WTO Disputes Over Environment, Health and Safety Issues

Thomas Sattler, Gabriele Ruoff and Thomas Bernauer*

Abstract

The quantitative literature on WTO dispute settlement explains dispute outcomes primarily through country characteristics and neglects policy-variables. The environmental policy literature tends to assume that disputes over inter-jurisdictional differences in environment, health and safety (EHS) regulation are particularly prone to escalation – without testing this assumption for the population of WTO disputes. We focus on these gaps by testing the claim that disputes over EHS regulation are more likely to escalate than disputes over other issues. The theoretical underpinning of this hypothesis is that, relative to non-EHS disputes, gradual concessions by the defendant to the plaintiff and compensation to domestic interest groups in the defendant country in exchange for international trade concessions are harder to achieve in EHS cases. We develop and test this hypothesis with data on 506 dyadic WTO trade disputes in 1995-2003, using selection models. The results show that, *ceteris paribus*, EHS disputes are less prone to escalation from the consultation to the Panel / Appellate Body level, but more prone to escalation into compliance disputes once they have reached the Panel / Appellate Body level. This finding suggests that escalation dynamics at different stages of the WTO dispute-settlement process differ across EHS- and non-EHS disputes and should be studied in greater detail in further research.

Correspondence: * Thomas Bernauer is a Professor of Political Science at ETH Zurich, Center for Comparative and International Studies (CIS), Weinbergstrasse 11, CH-8092 Zurich, Switzerland; Phone: +41-1-632-6466; Fax: +41-1-632-1289; bernauer@ir.gess.ethz.ch.

The authors are grateful to Rhena Forrer and Michael Cemerin for assistance in compiling the dataset, to Patrick Kuhn for advice in regard to the statistical procedures, and to Alasdair Young for comments on the coding of EHS cases.

1. Introduction

Global trade diplomacy in the GATT and, since 1995, the WTO has produced far-reaching cuts in tariffs and quotas. This has encouraged policy-makers to expand the focus of trade-liberalization efforts to non-tariff barriers to trade, such as domestic policies, regulation, and administrative practices that also affect the international flow of goods and services. Environmental, health and safety (EHS) issues have become particularly contentious in this context. Free-trade advocates often claim that EHS-regulation targeting products (e.g., packaging norms, food safety standards) is frequently designed so as to have protectionist effects – i.e., to impose disadvantages on foreign vis-à-vis domestic producers. They also claim that EHS regulations targeting production processes bias competition by affecting production costs and investment flows. Proponents of environmental and consumer protection, in turn, argue that EHS regulation has nothing to do with protectionism. They view regulatory differences between countries, even if they at times affect international flows of goods, services, and capital, as an expression of legitimate differences in social demands for protection of public health and the environment.

These opposing views clash almost regularly in the WTO. The WTO has sought to set legal norms in an effort to distinguish protectionist EHS regulation from non-protectionist EHS regulation. The key provisions to that end are included in the SPS- and TBT-Agreements of 1994 and Article XX of GATT. Nonetheless, trade disputes over the interpretation of these provisions – i.e., disputes over the lawfulness of particular domestic EHS regulations – arise almost regularly.

Political scientists have studied EHS disputes primarily by means of qualitative case studies. With reference to disputes over growth hormones, dolphins, turtles, asbestos, agricultural biotechnology, and other issues this literature tends to assume that WTO disputes over EHS issues are particularly explosive. In a recent article, Alasdair Young (2005) has challenged this assumption. He shows that EHS disputes in the WTO are much less frequent than often assumed. His analysis demonstrates that only very few of the many EHS issues raised in the SPS-committee of the WTO escalate into formal WTO disputes. Young's work is a very useful point of departure for the research reported in this paper. But it leaves us with two questions. First, it does not assess the *relative* propensity of EHS disputes (vis-à-vis non-EHS disputes) to escalate. This question can only be addressed if we study EHS disputes in the context of all WTO disputes. Second, Young's work does not provide an answer as to why (if so) EHS disputes are more (or perhaps less) prone to escalation than non-EHS disputes.

The quantitative literature on GATT/WTO disputes should, in principle, be able to provide answers to the two questions just raised because it usually studies the population of these disputes. But there are two gaps in this respect. First, the quantitative trade-disputes literature tends to focus on country-

characteristics as the main explanatory variables. It largely ignores policy-variables, i.e. the characteristics of disputed domestic policies or regulation. In particular, we do not know of any quantitative analysis that examines whether and why EHS problems are more (or less) prone to escalation in the WTO. Second, the dependent variable is usually defined in terms of concessions by the defendant, and not in terms of escalation per se. To the extent such studies focus on escalation they measure escalation as a two step process – from the consultation stage to the panel stage. They do not distinguish cases that are settled after a panel’s verdict from cases that escalate further into disputes over the implementation of panel verdicts.

Our paper seeks to fill these gaps. It puts forward a theoretical argument on why a specific policy-area in global trade relations (EHS issues) might be particularly conflictual. This argument holds that EHS disputes are riskier because gradual concessions by the defendant to the plaintiff and compensation by the defendant to domestic interest groups in exchange for consent to international concessions are harder to make in EHS than in non-EHS cases. We test this argument for 506 WTO dispute-dyads in 1995-2003. We also contribute to measuring escalation in a more nuanced way (three instead of two stages of escalation) and solving the associated methodological problems (selection effects).

The empirical results show that, *ceteris paribus*, EHS-disputes are less likely than other disputes to escalate from the consultation to the Panel / Appellate Body (AB) stage. This result contradicts the widely shared assumption in the qualitative case studies literature. However, we also find that EHS disputes, once they have reached the Panel / AB stage, are more likely to escalate into compliance disputes than non-EHS disputes. We derive several possibilities for further theoretical and empirical research from these results.

We begin with a review of the quantitative literature on GATT/WTO trade disputes, with a focus on policy-variables and EHS issues. The second part outlines a theoretical argument on why EHS disputes should be more prone to escalation than non-EHS disputes. The third part discusses variable definitions, the dataset, and the statistical procedure. The fourth part contains the empirical findings. We conclude by outlining options for further research.

2. Review of Existing Literature

The existing literature that is relevant to explaining trade dispute escalation concentrates on two areas that do not (yet) systematically connect. The trade policy literature explains state behavior (notably, liberalization vs. protectionist behavior) primarily with reference to variation in factors of production, asset specificity, and interest group behavior. The quantitative trade disputes literature focuses on explaining concessions by the defendant country. The explanatory variables are primarily country characteristics (e.g., extent of democracy, trade dependence, economic power). That is, the trade

policy literature emphasizes properties of economic policy areas, whereas the trade disputes literature emphasizes country characteristics. By concentrating on policy-variables in explaining trade dispute outcomes we connect these two areas of research to some extent. We pay particular attention to the quantitative literature because it is more directly relevant to this paper than the qualitative case studies literature.

Research on GATT and (from 1995) WTO trade disputes has centered on explaining the initiation of disputes, the choice of institutional mechanisms for conflict resolution, and dispute outcomes (notably, trade concessions by the defendant). Explanatory variables include the characteristics of countries involved (e.g., level of development, size of the economy, trade dependence, democracy, ideological orientation of government) and some strategic factors, such as retaliatory legal action and bandwagoning (e.g., Reinhardt 2000; Busch/Reinhardt 2002a; Garrett/Smith 2002; Davis 2003; Cemerin 2004; Leitner/Lester 2004; Dixon 2004; Martinez 2004; Widsten 2004).

We see primarily two gaps in the existing literature. First, the dependent variable is most often defined in terms of trade concessions by the defendant country (e.g., Busch/Reinhardt 2002a, 2002b; Garrett/Smith 2002). This definition of "success" produces an incomplete measure of escalation/de-escalation and the intensity of disputes. One might argue that the stability of the global free-trade regime hinges more on how the system copes with escalating trade disputes than on whether defendant countries ultimately make trade concessions in specific areas as a result of disputes. In fact, there are several cases in which the defendant made few or no concessions and the dispute did not escalate. The variables escalation and concessions by the defendant correlate only to some extent (see also Garrett/Smith 2002; Cemerin 2004).

Second, only very few pieces of research have paid attention to the characteristics of disputed policies or regulation. Some authors have used policy characteristics as control variables. Busch and Reinhardt (2002a, 2002c), for example, include "agricultural case", "discriminatory measure" and "sensitive case" (defined broadly as SPS cases or disputes over "cultural matters"). They observe (without theoretical underpinnings) that in their sample of GATT/WTO-Disputes in 1980-2000 "sensitive" cases were less frequently associated with concessions by the defendant. Similar results were obtained for 85 US-EU GATT /WTO disputes in 1960-2001. In the latter sample, concessions were 43% less likely (Busch/Reinhardt 2002a). These results suggest that EHS-cases might be more prone to escalation. However, there is a need to define and measure escalation and policy-variables and develop a theoretical argument on why escalation in some policy-areas is more (or less) likely. Some recent work has moved in this direction.

In explaining why only few WTO disputes escalate Garrett and Smith (2002) concentrate on three factors: strategic restraint of plaintiffs, primarily in cases where the defendant is unlikely to make concessions; strategically motivated

efforts by the Appellate Body to dampen conflicts, primarily when powerful states are involved; bargaining among plaintiff and defendant over the extent and timing of concessions. Garrett and Smith argue that these three mechanisms serve as safety valves for the most difficult WTO cases and thus reduce the likelihood of escalation. The empirical analysis concentrates on concessions by the defendant (and not escalation per se). Moreover, it seems difficult to reverse the three hypotheses in order to explain the likelihood of escalation (i.e., explaining when and why the three safety valves fail). Notably, Garrett and Smith do not explain why specific cases, which ex ante appeared prone to escalation, did not escalate (due to the effect of the three safety valves). They argue, e.g., that “doggedly pursuing sensitive, high-stakes cases...poses a significant threat to the future of the WTO”. Hence we should expect that the US-EU disputes escalate only when the anticipated damage to the WTO as a whole is small. It appears difficult to formulate and test ex ante predictions of this nature.

Garrett and Smith’s argument about strategic behavior of the Appellate Body may lead to the conclusion that disputes involving less powerful defendant countries are more likely to escalate because, in those cases, the Appellate Body is less likely to soften demands imposed on the defendant country. Empirical results (e.g. Busch/Reinhard 2002b) do not support such an argument. In respect to the third safety valve, bilateral bargaining over implementation of WTO verdicts, the two authors note: “Such breakdowns of bargaining could occur either because the actors did not accurately understand each other’s preferences and constraints or because domestic political considerations created an essentially empty set of acceptable outcomes in the international bargaining game.” They do not systematically test this argument. As to WTO disputes over bananas and foreign sales corporations they point to several conditions for escalation: (1) the dispute existed already in the GATT, (2) there already exists a WTO verdict in this area against the defendant country, (3) a treaty violation is obvious, but the prospects for concessions by the defendant country are poor, (4) the defendant country did not expect to be sued, (5) retaliation against legal action by the defendant in another area. Garrett and Smith’s argument also suggests that the Appellate Body will help in mitigating disputes when the legal foundation of a dispute is more open or unclear. Conversely, we should expect more escalation of disputes where the legal foundation is clearer. In view of several prominent disputes, such as hormones, biotechnology, dolphins, turtles, or asbestoes, this claim would obviously be problematic. In an event, it seems hard to come up with ex ante predictions that are based on whether or not the legal situation is clear or ambiguous.

As noted above quantitative studies of trade disputes connect only marginally to the standard trade policy literature. Recent work that focuses on interest groups and democracy is beginning to build such a bridge (albeit without reference to policy-variables). Davis (2003) argues that disputes involving strong interest groups in the defendant and the plaintiff country are more likely to escalate. It remains open, however, whether the strength of interest

groups can be measured reliably across disputes, countries, and over time. There are additional problems with conventional interest-group explanations as well. According to the Olsonian logic of collective action interest groups involving few actors and expecting concentrated costs (or benefits) are easier to organize, which is a necessary (albeit not sufficient) condition for political influence. Such groups may push the plaintiff country towards dispute escalation. One might also argue, however, that compensation of such groups (because they are smaller and easier to identify) by the plaintiff and/or defendant country should be easier to achieve. This might reduce the potential for escalation. Conversely, we often observe that in EHS disputes the contested policies of the defendant country are supported by a broad coalition of interest groups. Examples include the transatlantic disputes over growth hormones and agricultural biotechnology. At the same time, domestic pressure on the defendant country not to back down is strong. And it appears difficult if not impossible to compensate the opponents of de-escalation. In which manner the Olsonian logic comes into play seems to hinge on the properties of contested domestic policies. Our theoretical argument below will build on this assumption.

Some authors include democracy among their explanatory variables and, in some cases, relate this variable to arguments about interest groups. Yet, empirical findings have remained ambiguous. Guzman and Simmons (2002) claim that import-competing (in defendant states) and export-oriented producers (in plaintiff states) are able to exert greater political pressure on governments in democratic countries. Concessions are thus associated with higher transaction costs in democratic countries. This increases the probability of escalation of trade disputes in which democratic countries are involved. Delegation of decision-making authority to international institutions, such as the WTO, can be advantageous for democratic governments if it diverts political pressure on them in contested trade issues. In addition, as noted by Busch (2000) and others, democracies cherish the rule of law; hence they tend to have more confidence in legal mechanisms beyond the nation state and are more likely to seek decisions by WTO panels. Guzman and Simmons (2002) do not find empirical support for these hypotheses. Reinhardt (2000) and Busch (2000) assume that interest groups are easier to mobilize in democracies, that producers find it easier to organize than consumers, and that politicians in democracies are more sensitive to interest group pressure. They also assume that both import- and export-competing producers in potential plaintiff countries are interested in liberalization in a given target-country. This leads to the hypothesis that trade disputes among democratic dyads are more likely. Their empirical analysis shows that democratic countries participate more frequently in trade disputes, and that such participation makes dispute-settlement harder. While there appears to be some support for the monadic version of the hypothesis, Dixon (2003) finds no significant and stable effects for the dyadic version.

Explanatory variables at the center of the traditional trade policy literature, notably factor endowments and asset specificity, have thus far not been used

to explain trade disputes. In principle, one could argue that the probability of escalation increases if variation of factor prices and asset-specificity in a given dyad is such that it creates strong preferences for liberalization in one country and strong preferences against liberalization in the other country. Operationalizing these explanations in ways that avoid the revealed preferences problem is difficult and has not been undertaken to date.

A study by Guzman and Simmons (2002) is most closely related to the research presented in this paper. The two authors examine for the time-period 1995-2000 whether a specific policy-variable affects the probability of WTO disputes to escalate from the consultation to the panel stage. They hypothesize that WTO disputes over policies where concessions have an all-or-nothing character are more likely to escalate. The empirical evidence supports this hypothesis for democratic dyads. We build on this work by Guzman and Simmons and try to deal with some problems therein.

First, the definition of the explanatory variable in the Guzman and Simmons paper („continuous“ vs. „lumpy“ or „all or nothing“) is quite broad; the coding thus remains problematic. While their coding of SPS-related disputes is straightforward it is questionable whether or not product classifications or „absence of required laws“ can be modified continuously (similar to tariffs or subsidies). Two additional problems stem from the exclusion of eleven cases where coding as continuous or discontinuous appears to pose problems, and from the exclusion of cases where consultations were identified (by the WTO) as pending for more than three years. Second, the theoretical argument focuses on transfers (i.e., trade concessions and thus a transfer of wealth) from the defendant to the plaintiff country. We submit that an additional focus on transfers within the defendant country is important: escalation/non-escalation may also hinge on whether the defendant country is able to compensate (or pay off) domestic interest groups that bear the costs of concessions to the plaintiff country. Third, the dichotomous definition of the dependent variable does not take into account the fact that many disputes are taken to the panel stage, but that (from the perspective of the countries concerned) an important crossroads (whether to escalate or de-escalate) comes also after the Panel and/or AB stage. Our empirical analysis shows indeed that the probabilities of escalation from the consultation to the Panel/AB stage and from the latter to compliance disputes differ. It will also show that the determinants of escalation may not be the same across the three levels of escalation. As to escalation from the consultation stage to the Panel/AB level our empirical results differ substantially from those of Guzman and Simmons.

Our paper also contributes to the qualitative case study literature on EHS-disputes in GATT and the WTO. Most of the existing work at least implicitly assumes that EHS-disputes are particularly conflictual (e.g., DeSombre 2000; DeSombre/Barkin 2002; Young 2003; Bernauer 2003; Caduff 2004). Alasdair Young (2005) has shown, however, that of all complaints raised in the SPS committee of the WTO only a small part ends up in the formal WTO dispute settlement process – SPS cases are a large subset of what we define as EHS disputes. The limits of Young’s analysis are that he examines only SPS cases, without reference to probability distributions within the

population of WTO disputes. Moreover, he does not offer an explanation of why, if so, EHS disputes may escalate more often than non-EHS disputes.

In other words, our research contributes several new elements to the existing literature. We concentrate on an important policy-variable and connect this variable to a theoretical argument on concessions and compensation. We measure escalation in terms of three levels and estimate the effects of our main explanatory variable with selection models. Hence we connect two areas of research that rarely communicate: first, research on environmental and consumer protection policy, which tends to highlight the conflictual nature of EHS problems in global trade diplomacy, but does not test this argument within the population of WTO disputes; second, quantitative research on trade disputes, which has thus far largely ignored policy-variables and has focused on concessions by the defendant rather than escalation.

3. Theory

With the exception of Guzman and Simmons (2002) the quantitative trade disputes literature has largely ignored policy-variables. We build on this work and on interest group theories to develop a theoretical argument.¹

In the standard case a plaintiff country carries a dispute into the WTO when its government is pressured by domestic producers who believe they suffer from tariff- or non-tariff-barriers imposed by the defendant country. This pressure usually subsides when the defendant country reduces or eliminates the contested trade barriers or compensates the affected producers in the plaintiff country. Concessions by the defendant and hence de-escalation may materialize in at least three forms.

First, the defendant country can compensate the plaintiff country and its exporters through direct cash-payments. This instrument is extremely rare in global trade diplomacy (see Oye 1993). Second, the defendant country can offer non-cash compensation, for example in the form of trade concessions outside the disputed area. That is, it can, e.g., compensate the plaintiff country for dropping its complaints against import-restrictions on fish by liberalizing

¹ Unfortunately, the law&economics literature focusing on why some domestic disputes go to court whereas others are settled out of court does not offer much help in (empirically) understanding WTO dispute escalation. This literature focuses on information asymmetries and (overly) optimistic behavior of parties to a dispute. It assumes that under conditions of full information and no transactions costs all disputes would be settled out of court. Prima facie, it appears implausible that EHS problems are, systematically, characterized by greater information asymmetry or a greater extent of (overly) optimistic behavior of the countries involved. One possibility is that EHS cases are more complex than non-EHS cases – which might be associated with incomplete information, more information asymmetry, and problems of over-optimism. Yet, the complexity of WTO cases appears very difficult to code in a reliable manner. In addition, over-optimism is hardly a function of complexity alone. We face similar problems in operationalizing arguments from traditional trade policy literature that focus on factor prices and asset-specificity.

textile imports. This instrument is also rarely used, probably for two reasons. It is difficult to target compensations of this nature on one specific country because the most favored nation principle obliges WTO members to apply trade concessions offered to one other WTO country to all WTO countries. In addition, compensation of this kind will not alleviate domestic political pressure on the plaintiff government – exporters suffering from the trade barriers imposed by the defendant country would not benefit from concessions in another than the disputed area. Third, the defendant country can offer concessions in the disputed area itself, i.e., lower or eliminate the contested trade barriers. This is by far the most frequent form of concessions and thus the main instrument for achieving de-escalation in a dispute.

Assuming that de-escalation is usually achieved through concessions by the defendant country within the disputed policy-area it is important to understand the characteristics of disputed policies. Are there policy-areas in which disputes are harder (or easier) to solve and in which escalation is, therefore, more (or less) likely? In principle, properties of policy-areas are an open category that can lend itself to infinite differentiation of policy-characteristics. In this paper we focus on one particular category, the distinction of EHS and non-EHS disputes, because doing so connects to an important (and policy relevant) empirical debate, and because we can connect this distinction to a coherent theoretical argument on trade dispute escalation. We submit that EHS disputes are more likely to escalate for two reasons:

First, defendant countries often find it difficult to gradually reduce trade-restricting EHS regulation. Such regulation, and associated levels of EHS protection, tends to rest on firmly anchored social values of electorates that are widely regarded as non-exchangeable – i.e. electorates tend to be unwilling to accept reduced EHS protection in exchange for reduced trade action by the plaintiff country in the WTO. As proposed by Guzman and Simmons (2002), EHS issues tend to be regarded particularly by the defendant country as “all-or-nothing” issues. It is, for example, hard to imagine that the European Union will let some hormone treated beef into its market. Note that this argument is relative, not absolute. It is harder (but not impossible) to trade reduced EHS protection for reduced legal action by the plaintiff country than to carry out similar exchanges in non-EHS disputes. In other words, “fine-tuning” of tariffs or subsidies in areas such as textiles, steel, or agriculture is easier to achieve than “fine-tuning” in respect to EHS protection levels. This argument constitutes a straightforward application of the Coase-theorem: transaction costs in escalation-reducing exchanges are higher in EHS than in non-EHS disputes (see also Oye 1993).

EHS disputes suffer from yet another problem. Qualitative case study research on a range of EHS disputes strongly suggests that EHS regulation most often does not cater to narrowly defined protectionist interests – the latter suspicion is frequently voiced by critics of EHS regulation. It shows in fact that contested EHS regulation is often supported by large parts of the electorate and broad coalitions of organized interests. We submit that EHS

regulation, on average, is supported by larger parts of the electorate than more conventional trade-restricting measures, such as tariffs or quotas on textiles or steel. Bans on growth hormones or genetically modified food in Europe, for example, are supported by large coalitions that range from left-wing environmental activists to right-wing conservative farmers. Substantial concessions by the defendant to the plaintiff usually require compensation (by the defendant) to interest groups that have hitherto benefited from the trade-restricting measures. In addition to the problem that, in EHS-cases, those benefiting from EHS protection tend to regard such protection as non-exchangeable, the scope and scale of those benefiting from EHS protection poses difficulties. If EHS regulation was, as some critics claim, merely a tool to serve protectionist, rent-seeking interests, compensation would be easy (the number of actors would be small, and the problem of exchangeability per se would be minor). Yet, as noted above, this assumption does not hold. In other words, it appears easier to compensate a few dozen firms in the steel or textile industry, or to adjust agricultural policies (e.g., by shifting from product based subsidies to direct payments) than to compensate countless environmental and consumer protection groups, or even large parts of the electorate as such – provided exchangeability per se was acceptable.

The latter argument is a modified version of the logic of collective action. Consumers are harder to organize than producers. But if they succeed in obtaining preferred regulation – in some cases by forming coalitions with producers – those regulations are more resistant to international pressure (in this case emanating from WTO disputes) than conventional trade-barriers that often rest on regulatory capture by specific industries. It is easier to compensate the latter in gradual and more targeted ways.

We assume that the effects just outlined are stronger in democracies; that is, transactions costs in compensating domestic beneficiaries of trade-restricting regulation in the defendant country are likely to be higher in democracies. Greater transparency and public debate in democracies are also likely to play a role. Once democratic countries have, through costly signals, adopted a firm position in trade disputes the domestic political costs of backing down are likely to be higher.

The argument outlined in this section expands on the work by Guzman and Simmons (2002) in several respects. We reduce the main explanatory variable from the „continuous“ vs. „lumpy“ or „all-or-nothing“ definition to EHS cases. This definition suffers from one drawback: the connection between the theory and the empirical definition of the variable is less tight than in the Guzman and Simmons contribution. But the advantages are more clarity in empirical measurement, and tighter connection with an important empirical debate in global trade and the environmental policy literature. Second, we extend the theoretical argument from a Coasian country-to-country exchange to compensation problems among the defendant country and actors within the defendant country (problem of exchangeability per se, scale and scope of

compensation required). Third, we define escalation in terms of three levels and employ selection models.

4. Dataset, Definition of Variables, Statistical Method

In this section we describe the dataset, define the variables, and discuss the statistical method.

4.1 Dataset

The dataset used for this analysis comprises all trade disputes that entered the WTO system since its creation (1995) until the end of 2003. We thus include all disputes that were formally registered with a DS (dispute settlement) number (see WTO document WT/DS/OV/23). We do not include disputes that occurred in the GATT system (1948-1994). The institutional characteristics of dispute settlement differ very much between the GATT and the WTO. In contrast to the WTO, defendants were able to stall and/or prevent dispute settlement, whereas the WTO dispute settlement procedure operates almost like a court. It is, therefore, very difficult to code the degree of escalation of GATT disputes, particularly in regard to compliance disputes (the third level of escalation in our definition, see below).

To make sure that disputes have had a chance to escalate we restrict the dataset to the 1995-2003 period. Due to the time-lines specified by the WTO dispute settlement procedure, disputes that entered the WTO system, say in early 2004, could still escalate from the panel level to the compliance dispute level (see below for the empirical definition of escalation). Disputes that were initiated in late 2003 should, in our judgement, have had the chance to escalate into compliance disputes by mid-2005 (the latest information from the WTO we used in compiling the dataset). Yet, it is possible that some conflicts in our dataset that have currently not escalated may still do so in the future. For example: a dispute that entered the WTO in 2003 could be settled in 2005, but could flare up again in 2006 and proceed to the highest level of escalation. This problem exists irrespective of whether we study trade concessions (as most of the quantitative literature in this area does) or escalation. This problem could presumably be reduced if we cut the dataset back in time. But it would come at the expense of empirical observations. Particularly problematic in our case is the fact that a rather large number of EHS disputes entered the WTO in 2003. Because the overall number of EHS disputes is rather small to begin with we have opted to include 2003. However, we will test the stability of the results by excluding 2003.

In line with the approach taken in most quantitative studies on trade disputes we code disputes dyadically (e.g., Horn et al. 1999; Busch 2000; Busch/Reinhardt 2003; Bagwell et al. 2004).² We do so for several reasons. First, countries involved in a trade dispute can negotiate concessions bilaterally, even if the dispute involves more than one plaintiff. Plaintiff countries can ask for consultations or establishment of a panel at different points in time (Busch/Reinhardt 2003). Our dataset shows, in fact, that in

² Many WTO disputes involve more than one plaintiff country (5 on average; maximum of 19, minimum of 1), but all disputes involve only one defendant country each.

many cases only a subset of countries which initially asked for consultations escalates the dispute beyond that level. The same holds for escalation from the Panel / AB level to the compliance dispute level. Second, dyadic coding allows for the inclusion, as control variables, of country-characteristics. Third, dyadic coding assigns greater importance in the statistical analysis to disputes involving more plaintiff countries. The principal disadvantage is that the main explanatory variable in our analysis does not vary across dyads within a given dispute. We adopt the dyadic approach, however, because aggregating different (dyadic) levels of escalation and country-specific control variables to the country-group level (when there is more than one plaintiff in a dispute) would be rather dubious.

4.2 Definition of Variables

Table 1 shows the definitions of the two principal variables in our analysis. Definitions of the other variables are provided in the appendix.

Table 1: Definition of principal variables

<i>Name</i>	<i>Definition</i>	<i>Sources</i>
Escal	Escal is the dependent variable. It is coded as three levels of escalation (1, 2, 3). 1 = the dispute did not escalate beyond the consultation level in the WTO (the dispute was formally declared as resolved and/or did not result in the establishment of a panel); 2 = a panel was established and/or the Appellate Body of the WTO was activated; 3 = proceedings according to Articles 22.5 and/or 22.6 were initiated (compliance disputes). ³	WT/DS/OV/23 (April 7, 2005) and other dispute-related documents
EHS	EHS is the principal explanatory variable. It captures whether (1) or not (0) a WTO dispute involves EHS issues. We first assessed on the basis of WT/DS/OV/23 and other WTO documents whether one of the parties to a dispute invoked the SPS agreement or article XX of GATT. We then looked at the arguments of plaintiffs and the defendant to establish the extent to which EHS claims/arguments were voiced in the dispute. For the EHS variable, this coding rule was applied quite generously. Any dispute that involved some (even if few) arguments	WT/DS/OV/23 (April 7, 2005) and other dispute-related WTO documents

³ Further differentiation of escalation levels, e.g. separate coding of Panel and Appellate Body levels, does not appear meaningful. A very large part of Panel-cases moves on to the Appellate Body. Panels that were first established and then dissolved before completion of their work (notably, in the form of a panel report) or disappear from the WTO records were coded as level 2 – we code the highest escalation level a dispute has reached in its history. We code a dyad as level 3 only if the relevant plaintiff state has formally referred to Articles 21.6 and/or 22.6. Countries that ask for third party status at the consultation stage are not included in our data set since we include only dyads that were, formally, at least part of the escalation to the consultation level. Countries that first appeared as plaintiffs but then changed to third party status (e.g. when the dispute moved from the first to the second level) are not coded as being part of the dispute at the higher escalation level.

	by plaintiffs or defendant on EHS issues was coded as an EHS dispute.	
EHSn	EHSn operates with a narrower (hence the “n”) definition of EHS cases. We re-examined each case coded as EHS to establish to what extent EHS arguments were voiced by the parties to the dispute. We then excluded cases where EHS arguments played only a minor role. This coding rule was applied independently by two persons to the 71 dyads coded as EHS cases according to the EHS variable definition. In a few cases where the two coders arrived at different results they both re-examined the WTO documents and decided by consensus on the coding. The same coding procedure was also applied for the EHS and Escal variables. For reasons of transparency we list all Escal, EHS and EHSn codings in the appendix. The correlation between our EHS and the „continuous / discontinuous“ variable of Guzman and Simmons (2002) is -0.33 (for the EHSn variable it is -0.34). This difference probably stems from a substantially different definition of the variable and from the fact that our dataset includes many more cases.	WT/DS/OV/23 (April 7, 2005) and other dispute-related WTO documents

4.3 Statistical Method

In principle, we could use various statistical procedures to examine escalation and its determinants across different levels, notably, ordinal and sequential regressions. We prefer selection models for several reasons. Ordinal models (e.g., ordered probit models) work with the assumption that the influence of explanatory variables is the same across all levels of escalation.⁴ This implies, in our case, that an independent variable that increases the probability of escalation at the consultation level of the WTO procedure also increases the probability of escalation at the panel level. This assumption can easily be violated, since it is, theoretically, possible that, for example, the effect of an explanatory variable changes direction from one level of escalation to another. Indeed, we will observe such an effect in our empirical analysis.

One solution to this problem is to use conventional sequential regressions, that is, separate binary regressions for each level of escalation. This would allow observation of changes in effects of explanatory variables across levels of escalation. Yet, this approach is problematic because it may suffer from selection-bias and hence produce systematic deviations of estimators from

⁴ The ordinal regression model with J stages is equivalent to $J - 1$ binary regressions based on the assumption that the coefficient for the independent variables is identical across the different levels of escalation (the assumption of parallel regressions).

their proper values. The reason, in our case, is that only those disputes that have escalated to the second level can escalate to the third level. The estimation of escalation propensity from the second to the third level is thus not based on a random sample, but on a sample that systematically deviates from randomness.⁵

The recent international relations literature usually employs selection models that are based on a statistical procedure developed by Heckman (1976; 1979).⁶ Those estimators are only suitable, however, if there is at least one explanatory variable that influences the selection process but is not relevant to the outcome equation. This restriction is problematic if – as in our case – the same explanatory variables appear to be relevant to both escalation steps. We thus use a procedure developed by Sartori (2003), which allows for the same independent variables at both stages.

The critical assumption of the Sartori estimator is that the correlation of error terms of the first and second equation is one. This assumption may not fully hold in our case but is quite plausible. As Sartori argues, this should be the case if „(1) selection and the subsequent outcome of interest involve similar decisions or goals; (2) the decisions have the same causes; and (3) the decisions occur within a short time frame and/or are close to each other geographically” (Sartori 2003: 112). These conditions should, by and large, be met in WTO trade disputes. The aims of dispute parties – to protect domestic industries from foreign competitors and domestic consumers from EHS risks, or to pry open foreign markets for domestic firms – should not change substantially during any given trade dispute. Moreover, the time-frames of disputes are short enough to exclude fundamental changes in domestic and international economic structures to a degree that would change the interests and behaviors of the states involved.

5. Dispute Escalation in the WTO

Our data set includes 506 pairs of countries (dyads) that were involved in a WTO dispute in 1995-2003 (Table 2).⁷ 71 out of 506 dyadic disputes were EHS disputes, i.e. 14 percent.⁸ A large number of EHS disputes was launched in

⁵ This implies that the sample is censored at a specific point. This problem was first highlighted in the quantitative conflict studies literature (see Morrow 1989).

⁶ Van de Ven and Van Praag (1981) and Dubin and Rivers (1990) propose modified estimators, based on Heckman's work, if the dependent variable is dichotomous at the selection and the outcome stage. Such estimators have been applied, for example, by Reed (2000) and Lemke and Reed (2001) in conflict research.

⁷ The differences between our dataset and the dataset compiled by Guzman and Simmons (2002) in the time-period where the two datasets overlap (1995-2000) are as follows. Guzman/Simmons code dispute DS 57 as two dyads, we code it as one dyad. Even though the description of this dyad by Guzman/Simmons is identical, one dyad is coded as continuous, the other as non-continuous. The following cases are missing in the Guzman/Simmons dataset: 64, 77, 88, 102, 110, 123-125, 127-131, 133, 134, 137, 143-145, 147-150, 153, 154, 157-159, 167, 168, 171-174, 180, 182, 183, 185-187, 191, 196-201. DS 80 is coded in our dataset as a USA-EU dyad, in the Guzman/Simmons paper it is coded as a Belgium-USA dyad.

⁸ 39 out of 305 WTO dispute cases (some of them involve more than one dyad) are EHS cases, i.e. 13%.

2003 (31 out of 71, i.e., around 44%). This could bias our results, albeit primarily at the second to third level of escalation, for virtually all disputes launched in 2003 will, by mid 2005, have had a chance to escalate to level two (Panel/AB).

Table 2: Trade disputes, sorted by year of initiation

Initiation	Non-EHS (0)	EHS (1)	Non-EHSn (0)	EHSn (1)
1995	21	7	23	5
1996	37	9	38	8
1997	51	2	51	2
1998	37	6	42	1
1999	55	2	57	0
2000	40	2	41	1
2001	24	3	27	0
2002	128	9	129	8
2003	42	31	44	29
Totals	435	71	452	54
Share	86%	14%	89%	11%

A large share of EHS disputes in our dataset is about agricultural issues, notably quarantine and food safety issues. Note, for comparison, that a recently published WTO survey lists only seven trade-environment disputes (not dyads). In 17 out of 71 EHS dispute dyads the EHS component is rather small, vague or unclear. For example, in some cases the plaintiff country pointed out that the contested regulation in the defendant country could not be justified with EHS concerns (even though the defendant country did not explicitly mention such concerns) because this regulation would not be covered by the SPS Agreement or GATT Article XX. Excluding these 17 dyads the share of EHS disputes decreases from 14% to 11% (EHSn).

This data shows that EHS disputes are a relatively rare event in the WTO. But are they perhaps disproportionately prone to escalation? Figure 1 provides a provisional answer. It shows the unconditional probability distributions of EHS (EHSn) and non-EHS (non-EHSn) dispute dyads across escalation levels.

Figure 1: Probability distributions of dispute dyads across escalation levels

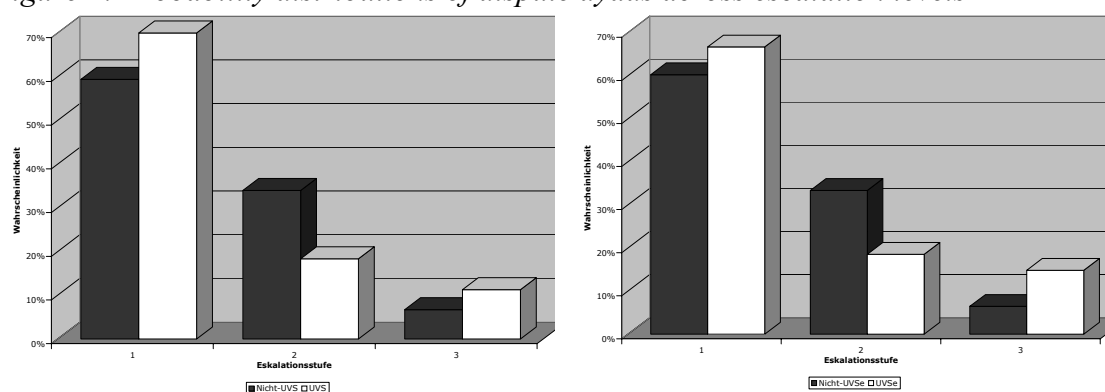


Figure 1 shows that EHS disputes are, in contradiction to our hypothesis, *less* likely than non-EHS disputes to escalate from the consultation to the Panel / AB level. Around 59% of non-EHS disputes remain at the consultation level, 70% of EHS disputes do so. The figure on the right shows that this difference is less pronounced for EHSn. A bivariate probit analysis (not reported in detail here) confirmed this visual finding. The estimated probability of escalation is 40% for non-EHS and 30% for EHS disputes.

The distribution from the Panel / AB to the compliance dispute level is very different from the distribution for the consultation to the Panel / AB level. It supports our hypothesis. EHS disputes escalate more often from the Panel / AB to the compliance dispute level. Almost 15% of EHSn disputes reach the third level, compared to 6% of non-EHSn disputes. This difference is smaller for the EHS coding. The conditional probability that a dispute escalates to the third level, given that it has reached the second level, is 38% (44%) for EHS (EHSn) disputes and 16% (16%) for non-EHS (non-EHSn) disputes. Bivariate probit estimates show that the according propensity of escalation rises from around 20% for non-EHS to around 40% for EHS disputes.

6. Multivariate Analysis

The empirical literature on trade disputes (e.g., Busch 2000; Busch/Reinhardt 2002a, 2003; Cemerin 2004; Dixon 2004; Widsten 2003, 2004; Garrett/Smith 2002) uses a variety of variables to explain the outcome of trade disputes (usually concessions by the defendant). We use the following variables as control variables:

- *Number of plaintiff countries:* We expect a negative effect of this variable on dispute escalation. The assumption is that political pressure on the defendant country to make concessions rises with the number of plaintiff countries.
- *Relative economic power of the plaintiff country:* We expect a negative effect of this variable on dispute escalation. The assumption is that defendant countries are more likely to back down in view of plaintiff countries that are economically more powerful and are thus able to employ larger “carrots and/or sticks” vis-à-vis the defendant. The new WTO dispute settlement system was designed to enhance the rule of law (as opposed to economic coercion). Whether this intention has dampened the effect of economic power on trade dispute outcomes is empirically open.
- *Trade dependence of the plaintiff on the defendant country:* We use an indicator that captures the trade dependence of the plaintiff on the defendant country relative to the trade dependence of the defendant on the plaintiff. We expect a negative effect. The assumption is that plaintiff states that are more dependent on trade with the defendant are less likely to escalate a trade dispute even if the defendant makes only small or no concessions.
- *Democracy:* The empirical literature on trade disputes has produced contradictory findings on the monadic and dyadic effects of democracy. For example, Busch (2000) observes that democratic dyads in the GATT system (until 1994) were better able to negotiate concessions in the consultation phase. But the probability of not reaching concessions was higher at higher stages of escalation. As noted in the theory section above we expect that the probability of escalation increases with the degree of democracy, and that the effect of the EHS variable could be stronger among democratic dyads.
- *Developing countries:* We expect a lower probability of escalation among developing country dyads because less resources are available to these countries for such purposes.

All explanatory variables used in the analysis are described in the Annex. Table 3 summarizes the descriptive statistics for all variables.

Table 3: Descriptive Statistics

<i>Variable</i>	<i>Mean</i>	<i>Standard Dev.</i>	<i>Minimum</i>	<i>Maximum</i>
Escalation 1st - 2nd level	0.389	0.488	0	1
Escalation 1st - 3rd level	0.462	0.628	0	2
EHS	0.140	0.347	0	1

EHSn	0.107	0.309	0	1
No. of plaintiffs	5.008	5.816	1	19
Rel. economic power	0.377	0.355	0.000	0.990
Trade dependence	0.031	0.109	-0.413	0.587
Democracy	19.777	2.321	6.481	21
Developing countries	0.132	0.339	0	1

6.1 Escalation from Consultation to the Panel / AB Level

Table 4 presents the results of our multivariate models.⁹ The upper part of the table shows the results for escalation from the first to the second level, i.e., the effects of independent variables on the probability of escalation of disputes from consultations to the Panel / AB level. The lower part of the table shows the results for second to third level escalation, i.e., escalation from the Panel / AB level to compliance disputes. Model (1) is the baseline model and contains only the control variables. In model (2) we add the EHS variable, in model (3) the EHSn variable.

The models including the EHS and EHSn variable perform better than the baseline model. We use the Akaike information criterion (AIK). In contrast to other measures based on the probability function it takes into account the number of independent variables in the model.¹⁰ Lower values of AIK indicate better model fit. The values for AIK in Table 4 show that model (2) performs best, followed by model (3).

⁹ These estimates are based on 502 dyads because bilateral trade data for Taiwan was not available.

¹⁰ We do not use correctly classified cases as a measure of model performance. This measure, which is usually based on a threshold value for correctly classified cases of 0.5, makes little sense if the distribution of the dependent variable across categories is uneven. As shown in Table 4, the number of escalating cases decreases strongly from consultations to the compliance dispute level. This would lead to the erroneous conclusion that more cases are correctly classified for the second equation (around 80%) than for the first equation (around 64%), even though model performance for the first equation (according to the AIK criterion) is in fact better. The reason for this difference is that, also because of reduced model performance at the second to third escalation level, the predicted probability of escalation exceeds 0.5 only in very few cases. Because only very few disputes in fact escalate to level three the measure for correctly classified cases over-rates model performance.

Table 4: Selection Models

	Model (1)	Model (2)	Model (3)
<i>Escalation from consultation to Panel / AB level</i>			
EHS		-0.413** (-2.37) -0.155	
EHSn			-0.237 (-1.22) -0.089
Number of plaintiffs	-0.083*** (-6.30) -0.031	-0.083*** (-6.36) -0.033	-0.082*** (-6.26) -0.031
Relative economic power of plaintiff	-0.565** (-2.59) -0.213	-0.606*** (-2.76) -0.227	-0.586*** (-2.68) -0.221
Trade dependence of plaintiff on defendant	-1.364* (-1.92) -0.513	-1.574** (-2.18) -0.591	-1.468** (-2.05) -0.552
Democracy	-0.028 (-1.05) -0.010	-0.024 (-0.91) -0.009	-0.025 (-0.96) -0.010
Developing countries	-0.822*** (-4.23) -0.309	-0.828*** (-4.26) -0.311	-0.831*** (-4.27) -0.312
Constant	0.975* (1.77)	1.017* (1.77)	0.993* (1.80)
<i>Escalation from Panel / AB level to compliance dispute</i>			
EHS		0.244 (1.08) 0.077	
EHSn			0.435* (1.82) 0.137
Number of plaintiffs	-0.060*** (-2.61) -0.019	-0.058** (-2.51) -0.018	-0.058** (-2.52) -0.018
Relative economic power of plaintiff	-0.563* (-1.66) -0.179	-0.558* (-1.67) -0.177	-0.548 (-1.63) -0.173
Trade dependence of plaintiff on defendant	-0.041 (-0.05) -0.013	0.058 (0.06) 0.018	0.088 (0.10) 0.028
Democracy	0.020 (0.44) 0.007	0.018 (0.39) 0.006	0.015 (0.32) 0.005
Developing countries	-0.748* (-1.82) -0.241	-0.753* (-1.82) -0.239	-0.691* (-1.65) -0.218
Constant	-1.359 (-1.42)	-1.355 (-1.40)	-1.317 (-1.38)
N	502	502	502
rho	1 (conditional)	1 (conditional)	1 (conditional)
Wald chi2	54.38	60.04	55.94
Pr > chi2	0.0000	0.0000	0.0000
AIK	809.163	800.986	803.722

*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level; z-values in brackets below coefficients; the third number in each cell indicates the estimated marginal effects, $dPr[\text{Panel}]/dx$ und $dPr[\text{Compliance}]/dx$, with all independent variables held at their mean; for dichotomous variables we calculate the change from 0 to 1 of the estimated probability of escalation.

The multivariate estimates confirm the visual impression obtained from Figure 1. The coefficients of the EHS and EHSn variables switch signs from the first-second to the second-third level of escalation. The upper part of table 4 shows that the EHS and EHSn variables tend to reduce the probability of escalation at the consultation level.¹¹ The estimated marginal effects (third number in each cell of the table) show that EHS disputes are more than 15% less likely to escalate from the first to second level than non-EHS disputes. The equivalent effect for the EHSn variable is 9%, but not significant.

With the exception of democracy the effects of the control variables are as expected. As can be seen in table 3 the democracy variable has questionable empirical properties to start with: the large majority of dyads score very high on the democracy variable; the mean of this variable (19,78) is very close to the maximum (21), the standard deviation is very small (2.32).¹² We will return to this variable below.

The number of plaintiff countries has a negative, significant effect: disputes with more plaintiffs are less likely to escalate at the first level. The probability of escalation decreases by 3% if the number of plaintiffs increases by one.¹³ It also decreases if plaintiffs are economically more powerful than defendants. The probability of dispute escalation among an economically very powerful plaintiff and a very weak defendant country is 20% lower than in disputes where the plaintiff is weak and the defendant strong. Disputes escalate less often the more dependent the plaintiff country is dependent in trade on the defendant country. These findings are interesting because the increased legalization of trade relations through the WTO's dispute settlement system was meant to reduce the role of power and coercion in global trade. Whether this has been the case is outside the scope of this paper because we do not include the GATT in our analysis. Finally, as expected, the probability of escalation decreases (by 26%) when both plaintiff and defendant country are developing countries.

6.2 Escalation from Panel / AB to Compliance Disputes

The lower part of table 4 shows the results for the second to third level of escalation, from the Panel / AB level to compliance disputes. Generally, we

¹¹ The upper part of table 4 (selection equation) is largely equivalent to a binary probit model for the first stage of escalation (from consultations to Panel / AB). The estimation of such a model with Huber-White standard errors (which are robust against heteroscedasticity) resulted in slightly higher z-values (significance levels).

¹² We included this variable in our model because it is considered relevant in much of the existing literature (Busch 2000; Guzman/Simmons 2002). The results do not change if this variable is excluded from the model. Other definitions of democratic dyads (e.g., Dixon 1994; Oneal/Russett 1997; Barbieri 2002) were tested as well. The results did not change.

¹³ The marginal effects reported in table 4 indicate how much the probability of escalation changes if a given independent variable varies in regard to its mean. Because models with binary variables are not linear the marginal effects for different values of the independent variables differ in size. The changes indicated in table 4 and the text are approximate values for the change in the explanatory variable by one unit in respect to its mean.

observe that the explanatory power of the models is lower for the first to second than for the second to third level of escalation.

As noted above, the effect of the EHS and EHSn variables changes direction: EHS disputes are more likely than non-EHS disputes to escalate. For the EHSn variable, dispute escalation is 13% more likely. For the EHS variable, the effect runs in the same direction but is statistically not significant. These results confirm that conventional, ordinal models would have been inappropriate for our analysis because they would not have shown that the effect of the EHS / EHSn variable changes across levels of escalation. We also conducted a Wald test (Brant 1990) to formally confirm that the effect of the EHS / EHSn variables is not identical across the two stages of escalation. This test (results not reported here) was positive.

The effects of the control variables in the outcome models run in the same direction as in the selection models. All control variables have a negative effect on the probability of escalation: this probability decreases the more plaintiffs are involved, the stronger (economically) the plaintiff country is, and if the dyad consists of developing countries. However, the marginal effects are somewhat weaker and less significant. Trade dependence and democracy have no significant effects in the outcome models.

We cannot test directly the argument that the effect of the EHS / EHSn variables is stronger among democratic dyads. The reason is insufficient variation in the democracy variable. The interaction term between the EHS / EHSn variables and a simple dyadic democracy variable (a democratic dyad is defined as a dyad where both countries score at least 7 on the Polity IV index) correlates very strongly (0.94) with the EHS / EHSn variable because most dyads in our dataset are democratic.¹⁴ To circumvent this problem we restricted the analysis to disputes among democratic dyads (those dyads where both countries score at least 7 on the Polity IV index, i.e., 432 observations). The effects of the EHS / EHSn variables in this censored sample are indeed stronger, but only slightly. The marginal effects of the EHS variable increase to -16.5% at the first to second escalation level and 9.2% at the second to third level. For the EHSn variable, they increase to -10.4% and 14.2% respectively.

We tested the robustness of these results in various ways. Table 5 shows that multicollinearity is not a problem. Of all independent variables only the number of plaintiff countries and relative economic power correlate substantially. This correlation is due to the fact that disputes with many plaintiffs are often directed against the EU or the USA. Because the EU and the USA are economically very powerful, all dyads including the EU or the USA as defendants score low on the relative economic power variable (defined as economic power of the plaintiff vis-à-vis the defendant). To evaluate this potential problem we divided the sample and ran the models without dyads

¹⁴ The technical problem is that the estimator of Sartori does not converge under these conditions. A probit analysis would be possible, but the results would not be very meaningful.

involving the EU or the USA as defendants. The effects of the EHS variable remain the same, but they become insignificant when we exclude the USA and the EU. The principal reason is that many EHS disputes involve the EU, the USA, or both; hence the number of EHS observations decreases dramatically. Finally, we restricted the analysis to disputes initiated before 2003. As noted above, a large share of EHS disputes were initiated in 2003, and these disputes may not yet have had a chance to escalate into compliance disputes. The results are largely similar as for the unrestricted sample. The effects of the EHS variables point in the same direction. However, the effects of both EHS variables is statistically insignificant in the selection models, but significant in the outcome models. The results for the second to third escalation level should be interpreted with great caution because the number of EHS / EHSn observations is very small when we exclude all disputes initiated in 2003.

Table 5: Correlations Among the Explanatory Variables

	EHS	EHSn	Trade dep.	Democr.	Devel. c.	No. of pl.	Rel. ec. p.
EHS	1.0000						
EHSn	0.8552	1.0000					
Trade dep.	-0.0794	-0.1017	1.0000				
Democr.	0.0636	0.0689	0.0951	1.0000			
Devel. c.	-0.0365	-0.0760	-0.1780	-0.1973	1.0000		
No. of pl.	-0.0349	0.0340	0.1258	-0.0745	-0.2224	1.0000	
Rel. ec. p.	-0.0358	-0.0928	-0.1969	-0.0224	0.2337	-0.5083	1.0000

Finally, on a speculative or prospective note, we recoded the dispute over agricultural biotechnology from level 2 to level 3. In October 2005, when our dataset was finalized, this dispute was still at the Panel / AB level. However, many experts noted that this dispute was likely to escalate further, similar to the dispute over bovine growth hormones. This recoding results in much stronger and more significant effects of the EHS / EHSn variables. We infer that, if we updated the dataset and ran the same models in 2-3 years from now the effects of the EHS variables would probably be stronger rather than weaker.

7. Conclusion

The results reported in this paper show that, contrary to a widely shared assumption in the case-study based literature, EHS disputes in the WTO escalate less frequently than non-EHS disputes from the consultation to the Panel / AB level. But they also show that, once EHS disputes have reached the Panel / AB level, they are more prone to escalation into compliance disputes. Additional research in this area should focus on several aspects that were not addressed in this paper.

First, one particular type of selection effect was not dealt with in this paper. As noted by Young (2005) for one sub-group of EHS disputes – disputes over SPS issues – only very few disputes of that nature escalate from the SPS committee to the consultation level in the first place (our analysis captures such disputes only from that level on). We suspect, moreover, that many trade disputes do not make it to the WTO at all, but are dealt with outside the WTO framework. In research on armed conflicts such selection effects are addressed by including the totality of country dyads for all years,

irrespective of whether a given dyad has experienced an armed conflict. A similar approach in our case could be to include all WTO member dyads for all years (since 1995) in which the countries concerned were members of the WTO. For each dyad and year we would have to include a fixed set of categories of trade relationships and code whether for any given dyad-year-trade area there was a trade dispute, how intense the trade dispute was, and whether it was an EHS or non-EHS dispute.

Second, the theory underlying the claim that EHS disputes are more likely to escalate is very simple. It could be extended based on arguments from the law & economics literature that deals with domestic legal action. Such extensions could come in verbal or formal (notably, game theoretic) form. Extensions and/or formalizations of the theory should, as a priority, shed light on why EHS disputes could be less likely than non-EHS disputes to escalate from consultations to the Panel / AB level but more likely to escalate from there to compliance disputes. One possibility is that the parties to a dispute are potentially aware of the high conflict potential of EHS disputes and therefore display strategic restraint at the consultation level. For that reason, they may carry to the Panel / AB only those disputes that they really care about. This, in turn, means that those EHS disputes that are in fact carried from consultations to the Panel / AB level are the most contentious ones and hence more likely to escalate.

Third, further research could incorporate trade concessions (the dependent variable in many quantitative studies on trade disputes) into the explanatory model. One of the open questions in this respect is whether concessions are a necessary or even sufficient condition for de-escalation or non-escalation. Finally, it would be interesting to include additional policy-variables that can be based on a coherent theoretical argument.

8. References

Achen, Christopher H. 1986: *The Statistical Analysis of Quasi-Experiments.* Berkeley CA, University of California Press.

Bagwell, Kyle/Mavroidis, Petros C./Staiger, Robert 2004: *The Case for Tradable Remedies in WTO Dispute Settlement* (World Bank Policy Research Working Paper 3314), Washington D.C.

Barbieri, Katherine 2002: *The Liberal Illusion: Does Trade Promote Peace?* Ann Arbor MI, The University of Michigan Press.

Bernauer, Thomas 2003: *Genes, Trade and Regulation: The Seeds of Conflict in Food Biotechnology,* Princeton, Princeton University Press

Brant, Rollin 1990: *Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression,* in: *Biometrics* 46: 1171-1178.

Busch, Marc L. 2000: *Democracy, Consultation, and the Paneling of Disputes Under GATT,* in: *Journal of Conflict Resolution* 44: 425-446.

Busch, Marc L./Reinhardt Eric 2001: Bargaining in the Shadow of the Law: Early Settlement in GATT/WTO Disputes, in: *Fordham International Law Journal* 24:158-172.

Busch, Marc L./Reinhardt, Eric 2002a: Testing International Trade Law: Empirical Studies of GATT/WTO Dispute Settlement, in: Kennedy, D. M./Southwick, J. D. (Hrsg.): *The Political Economy of International Trade Law: Essays in the Honor of Robert Hudec*, Cambridge, Cambridge University Press.

Busch, Marc L./Reinhardt, Eric 2002b. Transatlantic Trade Conflicts and GATT/WTO Dispute Settlement. (Conference on Dispute Prevention and Dispute Settlement in the Transatlantic Partnership, EUI/Robert Schumann Centre, Florenz, Italien, 3.-4. Mai 2002).

Busch, Marc L./Reinhardt Eric 2002c: Developing Countries and GATT/WTO Dispute Settlement, Manuskript.

Busch, Marc L./Reinhardt Eric 2003. Developing Countries and GATT/WTO Dispute Settlement, in: *Journal of World Trade* 37: 719-735.

Caduff, Ladina 2004: Vorsorge oder Risiko? Verbraucher- und umweltschutzpolitische Regulierung im europäisch-amerikanischen Vergleich: Eine politökonomische Analyse des Hormonstreits und der Elektronikschrott-Problematik, (Universität Zürich: Dissertation), Zürich.

Cemerin, Michael 2004: Institutioneller Wandel und Macht im Welthandelssystem: Die Intensität von Handelskonflikten im GATT-1947- und WTO-Streitschlichtungsverfahren, (Universität Zürich: Dissertation), Zürich.

Davis, Christina L. 2003: Setting the Negotiation Table: the Choice of Institutions for Trade Disputes (Annual Meeting of the American Political Science Association, Philadelphia, 28.-31. August 2003).

DeSombre, Elisabeth/Barkin, Samuel J. 2002: Turtles and Trade: The WTO's Acceptance of Environmental Trade Restrictions, in: *Global Environmental Politics* 2.

DeSombre, Elisabeth 2000: Domestic Sources of International Environmental Policy: Industry, Environmentalists and U.S. Power, Cambridge MA: MIT Press.

Dixon, Gregory C. 2004: Disputes for Votes: Institutional Variabion Among Democracies and Trade Dispute Propensity, (Annual Meeting of the International Studies Association, Montreal, 17.-20. März 2004).

Dixon, William J. 1994. Democracy and the Peaceful Settlement of International Conflict, in: *American Political Science Review* 88: 14-32.

Dubin, Jeffrey A./Rivers, Douglas 1990: Selection Bias in Linear Regression, Logit and Probit Models, in: *Sociological Methods and Research* 18: 360-390.

Garrett, Geoffrey/McCall Smith, James 2002: *The Politics of WTO Dispute Settlement*, Manuskript.

Guzman, Andrew/Simmons, Beth A. 2002: To Settle or Empanel? An Empirical Analysis of Litigation and Settlement at the World Trade Organization, in: *Journal of Legal Studies* 31: S205-S235.

Heckman, James J. 1976: The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependant Variables and A Simple Estimator for Such Models, in: *Annals of Economic and Social Measurement* 5: 475-492.

Heckman, James J. 1979: Sample Selection Bias as a Specification Error, in: *Econometrica* 47: 153-161.

Horn, Henrik/Mavroidis Petros C./Nordström Hakan 1999: Is the Use of the WTO Dispute Settlement System Biased? (CEPR Working Paper 2340).

Leitner, Kara/Lester, Simon 2004: WTO Dispute Settlement 1995-2003: A Statistical Analysis, in: *Journal of International Economic Law* 7:169-181.

Lemke, Douglas/Reed, William 2001: War and Rivalry Among Great Powers, in: *American Journal of Political Science* 45: 457-469.

Marshall, Monty G./Jagers, Keith/Gurr, Ted 2002: Polity IV Project. Political Regime Characteristics and Transitions, 1800-2003.

Morrow, James D. 1989: Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining, in: *American Journal of Political Science* 33: 941-972.

Neumayer, Eric 2004: The WTO and the Environment: Its Past Record is Better than Critics Believe, but the Future Outlook is Bleak, in: *Global Environmental Politics* 4:1-8.

Neyer, Jürgen 2005: External Effects of Multi-Level Governance – The EU and the US in International Trade Dispute Settlement, Manuskript, <http://www.soz.uni-frankfurt.de/neyer/>

Oneal, John R./Russett, Bruce 1997: The Liberal Peace: Interdependence, Democracy, and International Conflict, 1950-85, in: *Journal of Peace Research* 33: 11-28.

Reed, William 2000: A Unified Statistical Model on Conflict Onset and Escalation, in: American Journal of Political Science 44: 84-93.

Reinhardt, Eric 2000: Aggressive Multilateralism: The Determinants of GATT/WTO Dispute Initiation, 1948-1998. Manuskript.

Sartori, Anne E. 2003: An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions, in: Political Analysis 11: 111-138.

Van de Ven, Wynand P./Van Praag, Bernard 1981: The Demand for Deductibles in Private Health Insurance, in: Journal of Econometrics 17: 229-252.

Vogel, David 1986: National Styles of Regulation: Environmental Policy in Great Britain and the United States, Ithaca NY, Cornell University Press.

Widsten, Amy L. 2004: Credible Coalitions: Bargaining and Enforcement in International Trade Dispute Resolution (Annual Meeting of the American Political Science Association, Chicago 2.-5. September 2004)

Young, Alasdair 2003: Political Transfer and 'Trading Up': Transatlantic Trade in Genetically Modified Food and US Politics, in: World Politics 55: 357-84.

Zangl, Bernhard 2005: Das Entstehen internationaler Rechtsstaatlichkeit. Streitbeilegung in GATT und WTO. Frankfurt/M.: Campus.

Annex

Table A1: Definition of Variables

<i>Name</i>	<i>Description</i>	<i>Source</i>
Start	Year when the trade dispute was formally initiated in the WTO	WT/DS/OV/23 (April 7, 2005) and other WTO
Plaintiff country	Country that launched the dispute settlement procedure	WT/DS/OV/23 (April 7, 2005) and other WTO
Defendant country	Country accused of violating WTO law	WT/DS/OV/23 (April 7, 2005) and other WTO
DS number	Dispute settlement number assigned to the dispute by the WTO	WT/DS/OV/23 (April 7, 2005) and other WTO
No. of plaintiffs	Number of plaintiff countries involved in a dispute (as defined by the DS number)	WT/DS/OV/23 (April 7, 2005) and other WTO
PolityP	Democracy score of the plaintiff country (e.g., the EU scores 10 on this scale)	Polity IV project (Marshall et al. 2002)
PolityD	Democracy score of the plaintiff country (e.g., the EU scores 10 on this scale)	Polity IV Projekt (Marshall et al. 2002)
Jdemo	We draw on the conflict studies literature to define the level of democracy of a dyad (e.g. Lemke/Reed 2001). To obtain positive Policy IV scores we add 11 to PolityP and PolityD separately. We then multiply the two sums and draw the square root of this product. One missing value for Hong Kong in 1996 was filled in with a score of 10.	Based on PolityP and PolityD
ExportsP	Exports by the plaintiff to the defendant; for the EU we used the sum of all exports by EU member countries	IMF Directions of Trade Statistics CD Rom
ImportsP	Imports by the plaintiff from the defendant; for the EU we used the sum of all imports by EU member countries	IMF Directions of Trade Statistics CD Rom
ExportsD	Exports by the defendant to the plaintiff; for the EU we used the sum of all exports by EU member countries	IMF Directions of Trade Statistics CD Rom
ImportsD	Imports by the defendant from the plaintiff; for the EU we used the sum of all imports by EU member countries	IMF Directions of Trade Statistics CD Rom
Trade dependence	This variable measures to what extent the plaintiff and the defendant are equally/unequally dependent on trade with the other country. We use directed trade asymmetry. It is defined as trade dependence of the plaintiff minus trade dependence of the defendant. Trade dependence of the plaintiff is defined as exports by the plaintiff (ExportsP) plus imports from the defendant (ImportsP) divided by the GDP of the plaintiff (GDPp). Trade dependence of the defendant is defined as exports by the defendant (ExportsD) plus imports from the plaintiff (ImportsD) divided by the GDP of the defendant (GDPd). We had to exclude Taiwan because no reliable trade data was available.	Based on ExportsP, ImportsP, ExportsD, ImportsD
GDPp	GDP of the plaintiff country; for the EU we	IMF Directions of Trade

	used the sum of GDPs of EU member countries	Statistics CD Rom
GDPd	GDP of the defendant country; for the EU we used the sum of GDPs of EU member countries	IMF Directions of Trade Statistics CD Rom
Relative economic power	GDPp divided by the sum of GDPp and GDPd	Based on GDPp and GDPd
Developing countries	A dyad is defined as a developing countries dyad if neither plaintiff nor defendant belongs to the OECD	www.oecd.org