TRADE GAMES:  
THE WTO’S ROLE IN DISPUTES—THEORY AND PRACTICE

By Roberto A. Perrelli* and Earl L. Grinols*

August 2001

Abstract

For many, the crowning achievement of the World Trade Organization (WTO) is its improved dispute settlement mechanism. Viewing trade disputes as a game played by governments, however, suggests a political approach the study of the WTO in conformity with the new theory of endogenous protection and government trade interventions. Formalizing and analyzing the dispute game played by the infringing and complainant countries subject to the WTO process suggests that lowered costs of litigation should lead to an increase in infringements and more dispute litigations. This paper examines these theoretical conclusions in light of the history of US Section 301, General Agreement on Tariffs and Trade (GATT) and WTO trade disputes. We apply count data techniques to identify the monthly pattern for “births” of trade disputes and test the hypothesis of a WTO structural break. We also implement a survival analysis of case lifespans. The evidence supports the view that the WTO increased the monthly incidence of trade disputes, while shortening their lifespan in agreement with the political game model.

JEL Codes: C-25, C-41, C-72, F-13.

Key Words: Trade disputes, trade games, Poisson regression, survival analysis, USTR Section 301, WTO.
TRADE GAMES: THE WTO’s ROLE IN DISPUTES—THEORY AND PRACTICE

I. Introduction

During the last twenty-five years, the United States has participated in over 200 international trade disputes involving 42 countries. Although reasons to initiate and pursue cases sometimes appear to be solely pecuniary, these disputes are strongly influenced by economic and political circumstances. The politicization of trade is increasingly evident in the politics surrounding the World Trade Organization, the growth of the anti-globalist movement, and the trade and environmentalist movement. The new political theory of trade policy approach emphasizes modeling, formalization, and testing of these political determinants. The new theory of economic endogenous protection, for example, is part of this promising trend. The central feature of the new theory is the grounding of trade-related decisions, including those of government, in rational optimizing behavior and greater realism in describing government objectives and the strategic and economic environment. According to this approach, impediments and interferences to trade can be understood, not as the cumulation of the blackbox output of history, but as responses to modelizable structures and incentives facing government and her agents.

The growing recognition of the political and strategic nature of trade has elevated the status of the WTO. Its institutions are increasingly central to the inter-national economics of the 21st century. Scholars such as Jackson (2001) have commented extensively on the enlarged juridical basis for the WTO, of which trade dispute settlement is foremost. Even so, formalization of the political motivations of trade disputes and endogenous protectionism is still in its infancy. None of the papers dealing with the determinants of trade disputes of which we are aware, for example, uses a statistical or econometric approach, and we have not found any theoretically-based studies of trade disputes explaining the incidence or duration of cases, and the impact of the World Trade Organization. The salient features of each trade dispute may include diverse elements and one must be cautious in interpreting them. Nevertheless, a theoretical analysis incorporating unifying common features, coupled with a quantitative analysis can be particularly valuable to understanding the mechanics of modern trade disputes.

The present paper has two main goals. The first is to provide a starting game-theoretic model that explains country incentives when they initiate and litigate international trade disagreements. We treat trade disputes as a game involving countries $\alpha$, $\beta$, and a random flow of opportunities to infringe. When an infringement opportunity of given characteristics is presented to country $\beta$, it makes the decision whether to infringe its WTO obligations with respect to country $\alpha$, the complainant. The decision to infringe is based on the benefits to $\beta$ of the infringement, $\beta$’s beliefs about how $\alpha$ will react, and the costs of adjudication and settlement on the solution path anticipated for the case. Once an infringement has been discovered by the complainant country’s affected

---


3According to Jackson, well over half, “maybe two-thirds or three-quarters of the effective work of diplomats, missions and secretaries in the WTO system is now related in one way or another to the dispute settlement process, including consideration of potential suits.” Jackson (2001), p. 190.

4Park and Eggers (2000) provide a valuable compilation of trade dispute facts of the first 5 years of the WTO.
industry, the industry petitions \(\alpha\)'s government to protect its WTO rights. \(\alpha\)'s trade representative then decides whether to accept the case for processing (prosecution), and chooses its strategy regarding a subsequent negotiated settlement with the defendant and whether to proceed to WTO litigation if settlement cannot be reached.

We model the dispute process as a repeated game where the stage game is a finite extensive form noncooperative game with one-sided incomplete information. There is a role for complainant country \(\alpha\) to develop its reputation for strong pursuit of cases. Without a credible reputation for prosecuting cases, other countries will feel free to infringe \(\alpha\)'s WTO rights with impunity and negotiated settlements will never be reached because the defendant has no incentive to accept them. The implications of the model are derived for different beliefs about costs of litigating and the benefits of infringement, one or both of which may be WTO dependent.

The second goal of the paper is to provide a sophisticated empirical analysis of the patterns of trade disputes, seeing whether they provide a match to implications of the theoretical model. We examine three distinct trade dispute environments: Cases conducted under the USTR (United States Trade Representatives) Section 301 system which run from 1975 to 2000, cases treated under the GATT (General Agreement on Tariffs and Trade) involving the US as complainant or defendant, and cases treated under the WTO, 1995-2000, involving the US as complainant or defendant. The first two samples provide a longer panel of observations, but the third sample provides a larger set of recent cases and countries. Limiting attention to cases involving the US allows us to compare the USTR Section 301 panel and cases involving GATT and the WTO.

Using these samples, we characterize the monthly caseload, case arrival, and case duration of trade disputes from 1975 to 2000 using modern parametric, semi-parametric, and non-parametric models. We apply Markov regression techniques to describing the population of cases through time, and suggest a maximum likelihood estimator and two quasi-maximum likelihood estimators to account for overdispersion in the data relative to a Gaussian model. We then conduct a duration analysis of trade disputes, using a semi-parametric (Cox Proportional Hazard) format which we argue seems to represent the data well, followed by a variety of parametric models (Weibull, Exponential, Gompertz, Gamma, Log-Normal, and Log-Logistic regressions) which allow us to consider goodness of fit issues. We check the main properties of the models, and then compare the findings with Nelson-Aalen hazard estimates. Using those techniques, we test for the existence of structural breaks in the pattern of disputes after the WTO advent that match the predictions of the simple repeated game.

The remainder of the paper is organized as follows. Section I reviews the statistics and stylized facts of trade disputes for the period 1975-2000. Section II discusses the theory of trade disputes as a game played by two countries (the complainant and the defendant) each of which acts strategically to achieve its own objectives. We describe the strategies, parameters, and solution of this non-cooperative game in section III. We then perform several experiments to see the effect of lowered litigation costs on the outcome. We show that the model implies an increase in the number of infringements and cases as a result of the streamlined and more transparent WTO adjudication process. Section IV empirically models the population of dispute cases. However, population is determined by case births and lifespans. Section V, therefore, examines the births and lifespans of cases separately. The data tend to confirm the outcomes suggested by the model. Under WTO auspices there are more cases of shorter lifespan. This conclusion is less strong for the set of US Section 301 trade disputes than it is for GATT/WTO cases. Section VI summarizes and makes concluding remarks.

\(^5\)For autocorrelated count data we use Markov Poisson, negative binomial and normal estimators. Overdispersion refers to an estimated conditional variance higher than the estimated conditional mean.
II. Trade Disputes in Practice

The advent of the WTO raises interesting questions that relate primarily to its impact on the “population demographics” of trade disputes. These include the number of active cases (“population”), the number of cases initiated (“the birth rate”), and the duration of individual cases (“lifespans”). We selected data from three “families” of cases having in common an association with the US. We found that the number of such cases seems to have increased since WTO advent, and they tend to be shorter on average with smaller tails of extremely long-lived cases. The data sets we examined were US Section 301 disputes (123 cases, from July 1975 until February 2000), GATT disputes (71 cases, from September 1975 to October 1994), and WTO disputes (111 cases, from May 1995 through August 2000). Of 71 GATT disputes, 44 involved the US (23 as the complainant, 21 as the defendant), while 60 of the 111 WTO cases involved the US (34 as the complainant, and 26 as the defendant). All USTR Section 301 cases, of course, involved the US as a complainant. The data included cases if they involved the US either as complainant or as defendant. The data was collected from public sources (USTR (2000) and WTO (2000)), and contain information about when each case started and finished and the nature of the dispute. There are 19 CASES of overlap between the GATT and the Section 301 data, and 12 cases of overlap between the WTO and the Section 301 data. GATT and WTO data sets do not overlap because GATT was replaced by the WTO. Figure 1 displays the relationship between the different data sets.

A Population

The populations of GATT, WTO, and Section 301 USTR cases are plotted in Figure 2 using a cubic spline fitting curve to display a measure of the “average” caseload for each type of case. In addition to smoothing, the cubic spline allows for a non-linear time trend for each of the series.

---

Figure 1: Number of Trade Dispute Cases in Sample Showing Overlaps

![Figure 1: Number of Trade Dispute Cases in Sample Showing Overlaps](image_url)

---

6During its entire period of operation (1947-1994), the GATT system dealt with 102 international trade disputes. More extensive work on the topic might include the 31 GATT disputes initiated before 1975. However, the treatment of the earlier cases very likely differed from the later ones, and we felt that less would be learned from comparisons involving cases from so long ago.

7For Section 301 cases, there are two distinct sources of initiation: cases brought by private firms and cases brought by the US Government (so-called self-initiated cases). Grinols (1989) and Grinols and Perrelli (2000) suggest that political variables help explain the pattern of business-initiated versus government-initiated cases through time.
Sample months are shown on the horizontal axis, and a vertical line marks the first month of WTO operation.

The summary statistics for the series give a flavor of the distinctions between USTR Section 301, GATT, and WTO cases. For the USTR Section 301 cases, the monthly average number of open cases was 10.04, the median number was 10.00, the 80th percentile was 14.00, the standard deviation was 4.23, skewness 0.26, and kurtosis 2.70. Therefore, the series is almost mesokurtic and symmetric, with similar mean and median, although the Jarque-Bera statistic of 4.37 (with Pr($\chi^2$) = 0.11) shows that the series departs a little bit from the normal behavior. Moreover, the series is overdispersed (the variance of 17.92 is greater than the mean). For the GATT data the main statistics are: Mean 4.10, median 3, 80th percentile 7, standard deviation 2.88, skewness 0.88, and kurtosis 2.90. The GATT data is overdispersed and presents Jarque-Bera statistics of 31.08 (with Pr($\chi^2$) = 0.00). The main statistics for the monthly caseload of WTO cases are: Mean 17.02, median 18.00, 80th percentile 21, standard deviation 5.91, skewness -0.84, and kurtosis 2.82. The WTO series is slightly negatively skewed but quite mesokurtic, with mean and median nearly the same. However, like the Section 301 series, the WTO series is overdispersed (variance of 34.93). Furthermore, the departure from the normal is aggravated by the smaller WTO sample size (the WTO sample covers 56 months, while the USTR Section 301 sample covers 294 months and GATT 238 months). The Jarque-Bera statistic is 6.73 (with Pr($\chi^2$) = 0.03), and the null hypothesis of normality is easily rejected. As suggested by our theoretical model, WTO cases occur with greater frequency and the population of WTO cases is larger than USTR Section 301 and GATT cases, conditional on the respective time-windows of each panel.

B Births

We can also observe how the advent of the WTO affected the “birth rate” of international trade disputes. Figures 3 and 4 plot the number of case initiations by month. There is no substantial change in the monthly birth rate of USTR Section 301 panel cases after advent of the WTO, but there is a noticeable increase after the WTO advent (marked by the vertical line) in the number
Figure 3: Section 301 Births

Figure 4: GATT, WTO Births
of WTO births compared to GATT. Later we will examine the reasons for these differences, and estimate their significance.

\[ \begin{align*} \text{C Lifespans} \end{align*} \]

In addition to population and births, the lifespan of international trade disputes is important to understanding the efficiency of the dispute resolution mechanism. The raw data is described in Figure 5. A more exhaustive study will be presented in Section V.

The USTR Section 301 cases, consisting of the 92 cases that do not overlap with other data in the Venn diagram of Figure 1, had an average lifespan of 588 days, with a standard deviation of 577 days. The GATT data, consisting of 48 cases, had an average lifetime of 595 days, with a standard deviation of 440 days. The WTO data, composed of 51 cases\(^8\) had the shortest average lifespan, 544 days, with standard deviation of 290 days.

We also computed the median, skewness and kurtosis of the series. The USTR panel had a median of 378 days, skewness of 1.95 and kurtosis of 6.47. The GATT panel had a median of 470 days, skewness of 2.83, and kurtosis of 12.39. The WTO panel had a median of 562 days, skewness of 1.38, and kurtosis of 8.52. All series are clearly leptokurtic, and slightly asymmetric. Also, USTR and GATT series have a long upper tail. In terms of normality, the Jarque-Bera statistic for the USTR panel is 105.00 (with Pr(\(\chi^2\)) = 0.00), for the GATT panel is 240.90 (with Pr(\(\chi^2\)) = 0.00), and for the WTO panel is 81.24 (with Pr(\(\chi^2\)) = 0.00). Therefore, the null hypothesis of normality is rejected for all series at the minimum level of significance. Improved dispute settlement timetables suggest that negotiators wanted WTO cases to have a shorter lifespan than GATT or USTR cases. Further tests of the significance of this hypothesis will be provided in Section V. A comparison of the tails in Figure 5 suggests that the greatest proportion of extremely long-lived cases occurred among USTR cases, followed by GATT and WTO in that order. One explanation that can be ruled out is that the absence of long-lived WTO cases is due to the fact that the sample begins in

\[ \begin{align*} \text{WTO} \end{align*} \]

\[ \begin{align*} \text{GATT} \end{align*} \]

\[ \begin{align*} \text{USTR} \end{align*} \]

\[ \begin{align*} 0, 500, 1000, 1500, 2000, 2500, 3000, 3500 \end{align*} \]

\[ \begin{align*} 0.0000, 0.0005, 0.0010, 0.0015, 0.0020, 0.0025, 0.0030, 0.0035 \end{align*} \]

\[ \begin{align*} \text{\(\chi^2\)} \]
1995 and covers only five years. In fact, only six cases were censored by the termination of the data panel. Hence the shorter tail is due to improved procedures and the stricter timetables mandated by the WTO dispute settlement procedures.

III. Trade Disputes in Theory

From the economist’s perspective, the upsurge in WTO trade disputes did not happen by chance, but was the result of rationalizing decisions by participant countries about the opportunity to infringe and the choice to litigate using the WTO good offices. Both decisions can be described in terms of costs, benefits, and expectations about how the trading partner will respond. In this section, we describe the elements of the simplest repeated game model that could be used to evaluate the trade dispute decisions of trading countries. The decisions modeled include the decision to infringe, the decision to pursue a case for remedy, negotiation of a settlement, and the decision to establish a WTO panel and seek WTO adjudication.

A Complainant and Defendant Decisions

Presume that $\beta$ has infringed the WTO rights of country $\alpha$ and that an affected domestic industry has brought the matter to the attention of the trade officer of $\alpha$. In the United States this officer would be the US Trade Representative (USTR) and we will sometimes use this abbreviation generically when we talk of the complainant country $\alpha$. The USTR (synonymous with country $\alpha$), decides to accept or reject the case. There is an opportunity cost to assigning resources to the case. If the case is rejected, the USTR can use the resources for other pursuits.

According to WTO principles, negotiated resolution of disputes is the first objective. The length of time that a case takes to adjudicate, the resources needed to pursue the case, and the uncertainty of winning a case, are all elements that affect the desirability of making a settlement offer. The USTR therefore “negotiates” a settlement offer with the defendant country $\beta$ that is preferable to full-length panel adjudication for $\alpha$, which $\beta$ can accept or reject. We are not particularly interested in details of the bargaining process of this phase of the game. Such details would determine, for example, what fraction of the potential gain or surplus from settlement accrues to the complainant and what fraction to the defendant. Instead we make the streamlining assumption that the complainant makes a “take it or leave it” offer to the defendant. A take it or leave it offer is the optimal bargaining format from the complainant’s viewpoint and means that all of the surplus of an accepted offer accrues to the complainant. This feature of the model could be modified if desired, at the cost of greater, and for our purposes, unnecessary complication.

The defendant’s willingness to accept the offer depends on the defendant’s type: Some defendants have higher settlement costs than others (these could be psychic or political as well as real costs). The complainant knows the probability distribution of defendant settlement costs, but not the precise number, which is private knowledge of defendant. This feature is meant to capture the fact that in real life the complainant and defendant countries may have faced each other in similar disputes before. The complainant knows a lot about the defendant, but may not know for this particular dispute how important the issue is to the defendant. The complainant must therefore balance increasing the likelihood of acceptance of a less demanding offer against the worsened terms of the offer. The defendant, in turn, compares the offer to what it would receive if it refuses. If the offer is refused and the case is adjudicated further under WTO auspices, there is a cost to adjudication that is borne by both countries. The complainant has a probability of winning the WTO case, as does the defendant. We assume that both the complainant and defendant are rational
self-interested deterministic decisionmakers. Faced with the same circumstances in a repeated play of the game, they will make the same optimal decisions.

Figure 6 outlines the process just described. Payoffs on top apply to $\alpha$, payoffs below to $\beta$.

Stage 1: Decision to Infringe. In stage 1 cases arrive on a random basis in parallel streams for each set of case characteristics. $\beta$ must decide whether to use the opportunity the case arrival provides to infringe $\alpha$’s WTO rights. If no infringement is chosen, the payoffs to both players are zero (retain status quo). The decision to infringe initiates the remainder of the process.

Stage 2: Case Acceptance. The USTR (country $\alpha$) chooses whether to accept the industry case against $\beta$. The opportunity cost of resources employed by $\alpha$ if the case is accepted is $v$ and the direct welfare cost to $\alpha$ of the infringement is $B$. Country $\beta$ benefits from its infringement of $\alpha$’s WTO rights by amount $D$. (We do not assume that this is a zero sum game. Therefore, $B$ and $D$ do not have to be equal).\(^9\)

Stage 3: Negotiation. If a case is accepted, the USTR makes a take it or leave it settlement offer to $\beta$. The settlement offer is made with $\beta$’s willingness to accede taken into account. Defendant country $\beta$ accepts or rejects $\alpha$’s offer. The solution to the game determines the size of the offer and the likelihood that it is accepted in repeated plays.

---

\(^9\)For example, if $\beta$ employed a restrictive tariff against $\alpha$’s product there would be welfare costs to $\alpha$ and possible terms of trade gains to $\beta$. One’s gain would not necessarily equal the other’s loss.
Stage 4: **Litigation.** If the offer is rejected, the USTR decides whether to proceed with litigation under WTO auspices or to drop the case.

Stage 5: **Adjudication.** Assuming a case proceeds under the WTO dispute settlement mechanism, nature decides whether the complainant or defendant wins. The probabilities of a win or loss given the type of case are known to complainant and defendant.

It is helpful to refer to Figure 6 for the computation of player values at the different stages. The complainant country knows that it faces repeated plays of the game for each of the different types of cases and defendants. All information is known to both countries, except that the complainant does not know the defendant’s type. For simplicity, we summarize the steps of the solution in a series of propositions.

### IV. Formal Solution

This section solves the repeated extensive-form game of section II for its natural Nash subgame perfect equilibrium. Readers who wish, may skip directly to section F where a summary of the implications of the game are collected.

#### A The Adjudication Decision

We solve the game by backward induction, beginning with the stage 4 decision by $\alpha$ to proceed or drop the case against $\beta$. If $\alpha$ litigates, the WTO establishes a panel and the case proceeds to trial. Complainant pays adjudication costs $\gamma C$, where upper case C represents total adjudication costs under WTO auspices and $\gamma$ represents complainant’s share. Adjudication costs include the direct time and resource costs of trying a case but may also include psychic elements. Nature decides whether the complainant or the defendant wins the case. Complainant wins with probability $\varepsilon$, and defendant wins with probability $1 - \varepsilon$. The probability $\varepsilon$, depending on the type of case, the litigants, and the WTO’s panel effectiveness is known to both countries.

The value to $\alpha$ of dropping a case is $\sigma v - c_L$. $v$ is the value of country resources devoted to the case and $\sigma v$ stands for the share of those resources whose use will be obviated for fraction of the period $\sigma$ if the case is dropped. $-c_L$ represents the costs to complainant (possibly psychic) from dropping the case. For $\alpha$ to choose litigation, the potential current-period benefits, $\varepsilon b + (1 - \varepsilon)(-B - c_L) - \gamma C$, must exceed the value of dropping the case. Thus,

**Proposition 1. Complainant’s Choice to Litigate:** The condition for complainant $\alpha$ to choose litigation if the negotiated settlement is rejected is

$$\varepsilon(b + B + c_L) - \gamma C - B > \sigma v.$$  

#### B The Settlement Decisions

We now consider stage 3 decisions. $S$ stands for the value of the settlement between $\alpha$ and $\beta$. If the defendant accepts, the value of the decision to $\alpha$ is $S$. The value to the defendant is $-S - d_S$ where $S$ is the settlement and $d_S$ is the additional settlement cost to $\beta$. As before, cost $d_S$ can represent political, psychic, and/or real costs to the defendant of accepting the offer $S$. There is asymmetric information regarding settlement costs: $d_S$ is known by the defendant but not by the complainant. The complainant does know the distribution from which $d_S$ comes, however. Countries facing large political difficulties should they comply with the settlement offer, for example, would have high $d_S$. 


We assume that there is a random choice by Nature of $d_S$, e.g. high cost or low cost, whose \textit{ex ante} probability distribution is known by participants. The point outcome for player $i$'s cost is observed by player $i$ only. Mas-Colell et al. (1995) supplies a more didactic assessment of Bayesian games with a similar feature.

Defendants with high costs of settlement (high $d_S$) have harder-to-satisfy rejection bounds and will therefore reject even less demanding (lower $S$) settlement requests. If the defendant rejects the settlement then the value of this decision to the defendant is $\varepsilon(-d_L) + (1-\varepsilon)D - (1-\gamma)C$, the expected value from panel adjudication. This presumes that $\alpha$ chooses to proceed to litigation, a necessary assumption that we impose for now and take up below. $\beta$ will reject the settlement if the value of rejection exceeds the value of accepting, $-S - d_S$.

We assume that $\alpha$ is a rational complainant, balancing the size of the settlement it wants against the diminished likelihood of a larger demand being accepted. Having faced country $\beta$ in disputes before, $\alpha$, knows that the distribution from which $d_S$ is drawn affects how $\beta$ will respond to the settlement amount $S$. We model $d_S$ as being a binary random variable. The choice of a distribution does not change the nature of the model but does affect the specific magnitude. Percentage $p$ of the time the defendant has low settlement costs $d_S = \underline{d_S}$ and with percentage $1-p$ the defendant has high settlement costs $d_S = \bar{d_S}$.

We assume inequalities (2) and (3), which affirm that there is room for successful negotiation with low-cost defendants.

\[
(\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C) - \bar{d_S} < \varepsilon b + (1 - \varepsilon)(-B-c_L) - \gamma C \tag{2}
\]

\[
(\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - d_S) < (\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - d_S) - \varepsilon b + (1 - \varepsilon)(-B-c_L) - \gamma C \tag{3}
\]

The term in the middle is the least $S$ that $\alpha$ will accept; any lower amount would mean $\alpha$ would be better off going straight to adjudication. The term on the right (from (3)) is the most that $\beta$ would be willing to pay (going to adjudication would be preferred to paying an amount greater than this) if $\beta$ has low settlement costs ($d_S = \underline{d_S}$). The expression on the left is the most $\beta$ can afford to pay if its settlement costs are high ($d_S = \bar{d_S}$). That the term in the middle lies between the two means there is room for mutually beneficial settlement with low-cost defendants. If the middle term exceeds both other terms, there is no mutually acceptable settlement and the game goes straight to adjudication. If the middle term is lower than both other terms, all cases are settled without need of WTO panels.\(^{10}\)

How high should $\alpha$ set $S$? Since settling is better than litigating, the complainant’s alternatives are clear: Offering $S$ lower than $\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - \bar{d_S}$ will not improve the chances that the offer is accepted (low-$d_S$ defendants would be willing to pay the higher $\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - \bar{d_S}$) but \textit{will} reduce the payment from offers that are accepted. Demanding \textit{higher} payment than $\varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - \underline{d_S}$ means the settlement will never be accepted. Asking $S = \varepsilon d_L - (1-\varepsilon)D + (1-\gamma)C - \bar{d_S}$, however, will be accepted $p$ percent of the time, when $d_S$ is low.

The remaining $1-p$ percent of the time the offer will be rejected, the case will go to adjudication and earn (lower) value $\varepsilon b + (1 - \varepsilon)(-B-c_L) - \gamma C$. This is the best that $\alpha$ can do. The weighted average of the acceptance value and adjudication value is higher than the value generated by any other settlement. Hence, we have the following proposition.

\footnote{We presume that a credibility condition, discussed below, holds so that a nondegenerate equilibrium applies whereby $\alpha$ litigates cases if its settlement offer is rejected. If the credibility condition does not hold, the threat to litigate is futile (not believed), settlements are always rejected, and cases always dropped. This leads to an equilibrium where $\beta$ infringes and $\alpha$ does nothing but propose spurned settlements. While such an equilibrium is possible in the context of the model, it is uninteresting for our purposes.}

10}
Proposition 2: Negotiated Settlements.

(i) The defendant will reject settlement \( S > 0 \) if

\[
S > \varepsilon(d_L) - (1 - \varepsilon)D + (1 - \gamma)C - d_S. 
\]

(ii) The unique settlement request is

\[
S = \varepsilon(d_L) - (1 - \varepsilon)D + (1 - \gamma)C - d_S. 
\]

Notice that the rejection bound given by the right hand side of equation (5) can be negative. In that case, the defendant will reject any positive settlement. An implication of the model is that countries that have low settlement costs \( d_S \) will tend to pay more in settlement and will accept settlement more frequently, according to \( p \).

C  \( \alpha \)'s Decision to Accept a Case

We can now consider the stage 2 decision of the USTR to accept or reject a case. In conformity with the arrival of infringement opportunities to \( \alpha \) and \( \beta \)'s choice to infringe, cases before the USTR will arrive in different streams characterized by \((B,C,D,b,c_L,d_L,\varepsilon,\gamma,(p,d_S,d_S))\). The value of accepting such a case is

\[
\Pi \equiv p[\varepsilon(d_L) - (1 - \varepsilon)D + (1 - \gamma)C - d_S] + (1 - p)[\varepsilon b + (1 - \varepsilon)(-B - c_L) - \gamma C] - v. 
\]

where the first square-bracketed term is the value to \( \alpha \) if the settlement offer \( S \) is accepted. The second square-bracketed term is the value of WTO adjudication if settlement is rejected and the case is litigated. If the USTR rejects the case, the value is \(-B\), hence the condition for the USTR to accept the case is, \( \Pi \geq -B \), where \( \Pi \) is given by (6). All of the above is predicated on the belief by the defendant country that an accepted case will be pursued through litigation by the complainant if settlement is rejected.\(^{11}\) This belief must be credible in the repeated game setting. We turn now to this credibility or reputation constraint.

D  Reputation and the Credibility Constraint

It is in \( \alpha \)'s interest to establish a reputation for pursuing infringement of its WTO rights if negotiated settlement is rejected. To complete this part of the model we specify the beliefs of \( \beta \) regarding the out-of-equilibrium situation where \( \alpha \) fails to litigate when it is in its interest to do so. Recall that \( \alpha \) is a deterministic decisionmaker and this is known to \( \beta \). We assume, therefore, that \( \beta \) believes \( \alpha \)'s decision in this play of the stage game will be identical to its decision in the same circumstances in future plays. This belief sharpens the results of the model, highlighting the reputational effect and making the consequences to \( \alpha \) dramatic. \( \beta \) concludes that if \( \alpha \) does not litigate today, it will not litigate in identical circumstances in the future. \( \beta \) will therefore reject negotiated offers because it knows there will be no litigation consequences. This affects the probability of settlements being accepted (in this case the probability would drop to zero) and \( \alpha \)'s future welfare would be hurt. Maintaining its reputation for pursuing WTO cases when a credibility constraint showing self interest is satisfied, therefore, is important to the welfare of \( \alpha \). Given \( \beta \)'s beliefs and the “no trembling-hand condition”—i.e. the complainant chooses deterministically—the value to \( \alpha \) of litigation after a settlement is rejected is

\[
[\varepsilon b + (1 - \varepsilon)(-B - c_L) - \gamma C] + \sum_{t=1}^{\infty} \delta^t \Pi. 
\]

\( \delta \) is the discount factor. The first square-bracketed term is the value for the remainder of the period obtained by litigating and the

\(^{11}\) Otherwise the probability of \( \beta \) accepting a settlement will change.
second term is the value obtained in future periods from following α’s policy where δ is the discount factor and Π is given in (6). The second term therefore represents the value to complainant of its reputation as a country that pursues its WTO cases.

If α drops a case in mid-period after a rejected settlement, we let σ denote the share of value \( v \) recovered in the remainder of the period by complainant’s trade representative. The value to the country of dropping a case (hence revealing that it is not in its interest to pursue adjudication in cases of this type) is \( \sigma v - c_L \). Credibility requires that the value of adhering to the strategy exceeds the value of deviating. Thus,

**Proposition 3. Complainant Credibility:** The credibility condition is

\[
[\varepsilon b + (1 - \varepsilon)(-B - c_L) - \gamma C] + \sum_{t=1}^{\infty} \delta^t \Pi - [\sigma v - c_L] \geq 0
\]  

where \( \Pi \) is given in equation (6).

For future reference we denote the left hand side of (7) as the credibility constraint function \( CC(x) \) where \( x = (B, C, D, b, c_L, d_L, \varepsilon, \gamma, (p, d_S, d_S)) \).

**E The Decision to Infringe**

Sections A through D described the decisions of the complainant and defendant countries after an infringement of complainant’s WTO rights has occurred. Nondegenerate equilibrium cases are sometimes settled by negotiation and sometimes by costly litigation. \( \beta \), taking account of \( \alpha \)’s litigation strategy, uses the arrival of a case to infringe only if it is in its ex ante interest.

**Proposition 4. Infringement Condition:** \( \beta \) uses a case to infringe if and only if the infringement condition is satisfied,

\[
\varepsilon (-d_L) + (1 - \varepsilon)D - (1 - \gamma)C > 0.
\]

For future reference we denote the left hand side of (8) as the infringement condition function \( IC(x) \) where \( x = (B, C, D, b, c_L, d_L, \varepsilon, \gamma, (p, d_S, d_S)) \).

**F Trade Games: Summary**

We are now able to summarize the trade dispute game and its equilibrium strategies. The infringement process delivers infringement opportunities to \( \beta \) described by \( x = (B, C, D, b, \gamma, c_L, d_L, (p, \underline{d_S}, \bar{d}_S)) \). \( B \) is the lost welfare experienced by \( \alpha \) due to infringement. \( C \) is the total cost of adjudication, share \( \gamma \) OF which is borne by \( \alpha \), and \( D \) is the welfare gain to \( \beta \) from infringement. The benefit to the complainant of winning a panel decision is \( b \), costs to the defendant of losing a WTO panel decision \( (-d_L) \), and costs to the defendant of negotiated settlement \( (-d_S) \). The costs to the complainant of losing a case or dropping it after a negotiated offer is refused is \( (-c_L) \). These costs and benefits can represent psychological or political as well as real costs. Assuming the credibility condition (7) is satisfied, \( \alpha \)’s policy is:

- Accept a case if \( \Pi \geq -B \) where \( \Pi \) is given by (6).
- Ask negotiated settlement \( S = \varepsilon d_L - (1 - \varepsilon)D + (1 - \gamma)C - d_S \) if the case is accepted.
- Litigate if negotiated settlement is rejected.

This is matched by the defendant’s strategy:
If $D < 0$, do not infringe.
If $D > 0$ and the case is such that $\alpha$ will reject it for prosecution, infringe.
If $D > 0$ and the case is such that $\alpha$ will accept it for prosecution, infringe if and only if the infringement condition (8) is satisfied.\textsuperscript{12}
Accept $\alpha$’s negotiated settlement $S = \varepsilon d_L - (1 - \varepsilon)D + (1 - \gamma)C - d_S$ if $d_S = d_S$ and reject otherwise.

Rejected cases go to adjudication where $\beta$ receives $D$ if it wins and $-d_L$ if it loses. $\alpha$ receives $b$ if it wins and $-c_L$ if it loses.

G The Role of Costs on Infringements and Dispute Litigations

Formalizing the repeated game sharpens our understanding of the role that improved dispute settlement may have through lower litigation costs and bringing heightened visibility and greater world opinion to bear on trade disputes. An increase in the Infringement Constraint set, $IC = \{x \mid IC(x) > 0\}$, enlarges the set of cases that will be used as infringement vehicles by $\beta$ and an increase in the Credibility Condition set, $CC = \{x \mid CC(x) \geq 0\}$, enlarges the set of cases that will be litigated by complainant. If both regions grow there will be more infringements, and more WTO dispute cases. The effect of a change that increases one set and diminishes the other depends on which effect dominates—more infringements versus a smaller fraction that are litigated or vice versa. Litigation costs appear in the model as parameter $C$. The effect of greater visibility and legitimacy to the outcome of a trade dispute appear in the model as larger returns to complainant, $b$ if there is a win and higher cost, $c_L$ if there is a loss.

Table 1: Effect of WTO on the Infringement Constraint and the Credibility Condition

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Infringement Constraint</th>
<th>Credibility Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>$-(1 - \gamma) &lt; 0$</td>
<td>$-\gamma + \frac{\delta(1-p)\varepsilon}{1-\delta} &lt; 0^*$</td>
</tr>
<tr>
<td>$b$</td>
<td>0</td>
<td>$\frac{\delta(1-p)\varepsilon}{1-\delta} &gt; 0$</td>
</tr>
<tr>
<td>$c_L$</td>
<td>0</td>
<td>$1 - (1 - \varepsilon)[1 + \frac{\delta(1-p)}{1-\delta}] &gt; 0^{**}$</td>
</tr>
<tr>
<td>$d_L$</td>
<td>$-\varepsilon &lt; 0$</td>
<td>$\frac{\varepsilon p}{1-\delta} &gt; 0$</td>
</tr>
</tbody>
</table>

*For $\delta p$ small. **See text.

Table 1 summarizes the effects on the infringement constraint set and the credibility constraint set of a change in litigation costs $C$, benefits to the complainant of a win $b$, costs to the complainant of a loss $c_L$, and the costs to an infringer of a loss. According to the table, a reduction in adjudication costs, $C$, results in more infringements (enlarged set IC) and more litigations (enlarged set CC). Greater visibility and heightened world opinion (higher $b$ and $c_L$) increase the number of dispute litigations by enlarging the set of cases that the complainant credibly finds in its interest to litigate.

\textsuperscript{12}This is the value to $\beta$ of an opposed infringement, one which $\alpha$ responds to.
This assumes that the probability the infringer will be found guilty, $\varepsilon$ is relatively high.\(^{13}\) There is no effect of a change in $b$ or $c_L$ on the decision of $\beta$ to infringe, so the increase in cases results not from more infringements but from a greater number of infringements resulting in panel adjudication. The increase in costs to the infringer losing a case results in fewer infringements, but more litigations, so the net effect on trade dispute cases depends on which effect dominates. Since the main motivation for undertaking a trade infringement is the direct benefit to the infringer $D$, changes to the much smaller $d_L$ might be expected to have a less important impact on the number of cases. If this is true, the main impact of a reduction in $d_L$ is more cases operating through its effect on increasing the share of infringements that result in litigations.

V. The Population of Disputes

The theoretical model predicts that improved dispute settlement and lowered costs of litigation should encourage more infractions and more litigations. A comparison of the number of active GATT versus WTO trade disputes over time seems to bear this out, although a comparison of USTR cases before and after the advent of the WTO is less clear. The econometric modeling of population that we use to reach this conclusion is described next. The following section decomposes population changes into its two determinants, explaining the effect of the WTO on births and lifespans separately.

A Markov Estimation

The number of active trade disputes at any point in time equals the number of new cases birthed at the start of the period, a count data phenomenon, plus the number of cases alive in the previous period that continue into the present, a time series phenomenon. An unmodified count data model, therefore, such as the Poisson that presumes independence of observations, would be inappropriate as a statistical descriptor, as would a time series model that doesn’t address the birth feature. Since population satisfies the Markov property that the conditional distribution of population in period $t$ depends solely on information available in the previous period, the solution proposed to model the number of trade dispute cases is Markov regression.\(^{14}\) Ideally we would predict the number of cases in period $t+1$ by estimating for each dispute $i$ at time $t$ probability $p_{it}$ of it terminating (dying) in that period and probability $1 - p_{it}$ of surviving into the next, plus estimate the probability of new disputes arising (births) based on characteristics of period $t$. Our data set, while not small by many standards, requires some simplification for meaningful estimates. We employ a first-order conditional Markov model, where all information necessary to infer the size of the current population is included in the first-order lagged dependent variable, plus exogenous variables. Other information from the past is not relevant. Given the persistence through time displayed by the number of trade disputes, this approach seems to be the most appropriate. Some assumption must be made about the distribution for the transition probabilities. Our choice is “observation-driven.” In addition to the normal we assume that the conditional distribution of the

\(^{13}\)If $\varepsilon = 1$ or is close to one, $\frac{\partial CC}{\partial c_L} > 0$. It is easier for the complainant to credibly show that it is in its interest to carry cases to adjudication because to drop a case leads to higher costs. The second term in line three of Table 1 (multiplied by $(1 - \varepsilon)$) represents the effect of losing a case once it has gone to a WTO panel. It is important only if there is sufficient chance that the case will be lost in panel review. If that risk is great, of course, then the complainant is less likely to bring cases.

\(^{14}\)See Cameron & Trivedi (1998).
series is Poisson or negative binomial (Cox (1981)). To check the Poisson assumption, we provide tests for Poissoness.

**B Estimation Issues**

**Forms.** The Poisson distribution gives the probability of observing \(y\) events in a unit time interval.

The density of the Poisson distribution conditional on the observation set \(x\) is

\[
f(y|x, \beta) = \frac{e^{-\lambda(x, \beta)} \lambda(x, \beta)^y}{y!}
\]

where \(\lambda\) is the distribution parameter, \(x\) is the set of explanatory covariates, and \(\beta\) is vector of regressors. The maximum likelihood estimator (MLE) for the Poisson distribution is calculated by maximizing the log-likelihood function:

\[
\ell(\beta) = \sum_{i=1}^{N} \{y_i \ln \lambda(x_i, \beta) - \lambda(x_i, \beta) - \ln(y_i!))\}
\]

where \(y_i\) is the expected number of events in a unit time interval, and \(x_i\) are the explanatory variables. Under the Poisson distribution, the conditional mean and variance are given by the trivial estimate

\[
\hat{\lambda} = \exp(\hat{\beta} x).
\]

In spite of its fine properties and natural interpretation, “Poissoness” implies that the conditional mean equals the conditional variance, a property that data frequently fails to match, especially in samples that exhibit overdispersion. Among other authors, Koenker (1988) interprets this phenomenon as representing an inherent variability in the Poisson parameter \(\lambda\) (the expected number of events per unit time) around its hypothesized log-linear form.\(^{15}\)

Following the current literature, we use negative-binomial quasi-maximum likelihood estimators (QMLE) to correct for overdispersion. Given data where \(\text{var}(y|x_i)\) is not orthogonal to \(x_i\), nonlinear least squares estimators will lack efficiency properties for estimating conditional means, while generalized linear models calculate the conditional mean and the conditional variance of \(y\) without assuming any particular distribution (Wooldridge, 1997). The use of QMLE relaxes the Poisson restriction and provides consistent estimators even when the conditional distribution of \(y\) is not known. The density of the negative binomial distribution conditional on observations \(x\) is

\[
f(y|x, \beta) = \binom{n + y - 1}{y} (1 - \theta(x, \beta))^y \theta(x, \beta)^n
\]

where \(y\) is the expected number of failures before \(n\) successes, \(n\) is the number of successes in a sequence of binomial trials, \(\theta\) is the probability of success at each trial, and \(x\) are covariates that determines \(\theta\).\(^{16}\) The log-likelihood for the negative binomial is

\[
\ell(\beta, \theta) = \sum_{i=1}^{N} \{y_i \ln(\lambda_i \theta^2) - (y_i + \theta^2) \ln(1 + \lambda_i \theta^2) + \ln \Phi(y_i + \frac{1}{\theta^2}) - \ln(y_i!)\}
\]

\(^{15}\)According to Cameron & Trivedi (1998), we may assume Poisson fit if a relationship between the parameter \(\lambda\) and the exogenous covariates \(x_i\), such as \(\ln \lambda = \beta x_i\), is parametrically exact, and does not comprise any endogenous term. If the relationship above is stochastic, e.g. \(\ln \lambda = \beta x_i + \varepsilon\) we should assume an alternate mixed-Poisson model such as the negative binomial.

\(^{16}\)See Bickel & Doksum (1977) for a better discussion of probability models.
where \( y_i \) is the expected number of events (failures) before the \( n \)th success, \( \lambda_i \) is the distribution parameter representing the mean, \( \theta^2 \) is the overdispersion-parameter, and \( \Phi \) is the distribution function of such parameter.

Finally, using the normal distribution as the baseline model, we also estimated a normal QMLE. The normal distribution as applied here likewise gives the probability of observing \( y \) events in a unit time interval. The density of the normal distribution conditional on observations \( x \) is

\[
f(y|x, \beta) = \frac{1}{\sqrt{2\pi \sigma}} \exp\left(\frac{1}{2\sigma^2}(y - \lambda(x, \beta))^2\right)
\]

where \( y \) is the expected number of events, \( \lambda \) is the conditional mean, \( \sigma \) is the conditional variance, and \( x \) is the set of covariates that explain \( \lambda \). The normal QMLE has the following log-likelihood

\[
\ell(\beta) = \sum_{i=1}^{N} \left\{ -\frac{1}{2\sigma^2}(y_i - \lambda_i)^2 - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi \right\}.
\]

where \( y_i, \lambda_i, \) and \( \sigma \) have the same meaning as in the former equations.

Despite the usefulness of this last model for comparison purposes, as \( y \) is not normally distributed, consistent estimates are only provided if the error terms \( \varepsilon_i \) are homoskedastic and the parameter \( \lambda_i \) is correctly specified. Pagan & Pak (1991), however, recall that there is an intrinsic heteroskedasticity in count data models; this rules out the possibility of obtaining preferred estimates through this model.\(^{17}\)

**Discrete-valued Time Series.** A property of our sample is the persistency of the discrete dependent variable. Besides this persistency, the series is discrete-valued, overdispersed, and we are dealing with modest samples. Conventional time series models do not suit those cases. One feasible solution is proposed by Zegger and Qaqish (1988) and Cameron and Trivedi (1998) using a two-step count regression. The first step controls for the zeros in the lagged dependent variable, using a location-transformation by adding a constant term \( c^* \in (0,1) \) to each element of the series, \( y_{t-1}' = y_{t-1} + c^*. \)\(^{18}\) The second step includes the transformed lagged dependent variable as a multiplicative term in the first-order Markov model, and estimates the regression:

\[
\ln(y_t) = \alpha + x_t'\beta + \gamma \ln(y_{t-1}') + \varepsilon_t
\]

which is our final model.

**Robustness.** To obtain robust standard errors, allowing for the possibility of overdispersion, we adopted the popular Generalized Linear Model (GLM) covariance matrix, \( Var_{GLM}(\hat{\beta}) = \hat{\sigma}^2 Var_{ML}(\hat{\beta}) \) where \( \hat{\sigma}^2 = \frac{1}{N-K} \sum_{i=1}^{N} \frac{(y_i - \hat{\beta})^2}{\sqrt{Var(\varepsilon_i)}}. \)

**Model Evaluation.** For model evaluation we referenced the following statistics: adjusted-\( R^2 \), log-likelihood, likelihood ratio (LR), the Akaike criterion (AIC), and the Schwarz criterion

\(^{17}\)We also considered negative binomial and exponential QMLE, but tests indicated that they were no better than the models presented in the paper.\(^{18}\)To find \( c^* \), the authors suggest regressing \( \ln(y_t) = \alpha + x_t'\beta + \gamma \ln(y_{t-1}^*') + (\gamma \ln c) d_t + \varepsilon_t \) where \( y_{t-1}^* = y_{t-1} \), \( d_t = 0 \) if \( y_{t-1} > 0 \), or \( y_{t-1}^* = d_t = 1 \) if \( y_{t-1} = 0 \). The optimal term would be \( c^* = \exp(\gamma \ln c)/\gamma \). In an application to U.S. strike data presented by Keenan (1985) and re-treated in Jaggia (1991), however, Cameron and Trivedi (1998) use an arbitrary value of \( c^* = 0.5 \) and find little difference in the estimates. Here we also USE 0.5.
Generally, models with large log-likelihood values provide lower bias, and a best fitting ability. Models with small AIC or SBC criterion present a better specification, based on the Bayesian prior about the existence of a “true” model. For sample sizes greater than 8, the SBC criterion attributes a larger penalty for each covariate added to the model than the AIC criterion does. According to Gourieroux & Monfort (1995), the SBC criterion has the advantage of being consistent. Thus, for large samples SBC provides a more parsimonious decision rule about covariate selection and model specification. Finally, for the special case of the Poisson distribution, we provide another useful criterion of econometric modelling, the Goodness-of-Fit test, which identifies inappropriate Poisson specifications when the test statistic is significant. To check for the presence of autocorrelation, we applied residual-based tests.

### C Results

The techniques presented above were used to model the number of trade disputes. The estimation results are presented in Table 2. The goodness-of-fit test for the Poisson model shows that the null hypothesis of “Poissoness” cannot be rejected at the 5% level in both panels. We also note that all regressions display an extremely high coefficient of determination (around 97% in the GATT/WTO panel, and 92% in the USTR panel). All regressions pass the likelihood-ratio test for significance. The normal models, as is usual, generate higher log-likelihood and consequently smaller AIC and SBC values, followed by the Poisson and the negative binomial models, respectively. But, as explained earlier, the normal model is inconsistent in the presence of heteroskedasticity, and Poisson and negative binomial models should be considered instead.

The estimates in Table 2 indicate that the advent of the WTO had a positive and significant impact on increasing the population of GATT/WTO cases by 25% in the Poisson regression, 33% in the negative-binomial regression, and 26% in the normal regression. The Uruguay Round contributed to increasing the population of trade disputes by 16%, 19%, and 16%, respectively in the Poisson, negative binomial, and normal models.

On the other hand, neither the WTO nor the Uruguay Round appear significantly in explaining the population of USTR Section 301 disputes. This result is a puzzle, given our expectations about the impact of WTO advent on the population of cases. A careful review of the data suggests one possible resolution of this puzzle. The Venn diagram provided in Figure 1 shows that the proportion of USTR Section 301 cases that are also GATT (19%) is much lower that the proportion of USTR Section 301 cases that are also WTO cases (55%). Of the 102 USTR Section 301 cases initiated up to December 1994, 19 were also GATT cases. In the other hand, of the 22 USTR Section 301 cases opened after December 1994, 12 were also WTO cases. This result suggests that the USTR is using the WTO dispute settlement mechanism more frequently than it had used the GATT apparatus. For this reason, a greater number of USTR Section 301 cases are also WTO cases nowadays, and there is a significant probability that a potential new USTR Section 301 case is converted into a WTO case before its initiation.

---

19Recall that \( \text{AIC} = -2\ln(l) + 2(k/n) \), while \( \text{SBC} = -2\ln(l) + k\ln(n/n) \), where \( l \) is the value of the log-likelihood, \( n \) is the sample size, and \( k \) is the number of parameters included in the model.

20Recall that the equations include the presence of first-order lagged dependent variables.

21The marginal contribution of the WTO, given the other covariates, is measured by the derivative of \( \ln(y_t) \) with respect to the WTO covariate. For example, \( \frac{d\ln(y_t)}{dWTO} = 0.2528 \) in the Poisson model.

22We ran a simple \( \chi^2 \) test for the difference of two proportions, and at 1% level of significance we can reject the null hypothesis that they are the same.
Table 2: Impact of WTO on the Population of Trade Disputes

<table>
<thead>
<tr>
<th>Panel Model</th>
<th>GATT/WTO Cases</th>
<th>USTR Section 301 Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poisson</td>
<td>N-Bin.</td>
</tr>
<tr>
<td>Ln(y*_t-1)</td>
<td>0.9514</td>
<td>0.9244</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.0317)***</td>
<td>(0.038)***</td>
</tr>
<tr>
<td>WTO</td>
<td>0.2528</td>
<td>0.3328</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.0631)***</td>
<td>(0.079)***</td>
</tr>
<tr>
<td>Uruguay Round</td>
<td>0.1556</td>
<td>0.1902</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.0440)***</td>
<td>(0.058)***</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.1236</td>
<td>-0.1085</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.0467)***</td>
<td>(0.052)***</td>
</tr>
</tbody>
</table>

| Goodness-of-fit test | 66.12 | 39.39 |
| Adjusted-R² | 0.97 | 0.97 | 0.97 |
| Log-likelihood | -519.31 | -769.20 | -411.58 |
| LR (\_\_\_) | 1430.08 | 194.37 | 11126.53 |
| LR > \chi²(6) | 498.09 | 50.40 | 4785.81 |
| AIC           | 3.57  | 5.28  | 2.84  |
| SBC           | 3.62  | 5.33  | 2.89  |
| Sample Size   | 293   | 293   | 293   |

***Significant at 1%. **Significant at 5%. *Significant at 10%.

VI. Dispute Births and Lifespans

Population change is the result of both births and deaths, but does not provide separate information about either. Knowledge about the births and deaths of trade disputes—or, equivalently, births and lifespans—therefore, provides additional insight into how the WTO may have affected trade disputes.

A Births: A Count Data Analysis

Using the same data as in the population study, we conducted a pure count data analysis of the birth of trade disputes. Instead of the number of open cases as the dependent variable, the number of new cases per month was the dependent variable. The inter-temporal WTO indicator variable was the only covariate of the model. This choice was made after testing for the significance of the other covariates that we included in the population study, and finding insignificant explanatory power for births. Retaining the WTO covariate provided results on the role of the WTO.

We applied three different models to births: the Poisson, the Negative Binomial, and the Normal. According to the theoretical model of Section III, lower costs of adjudication made possible by the WTO should lead to an increase in cases. This is what we found for WTO cases, but not for USTR Section 301 cases. The results are shown in Table 3 below.

From Table 3 we see that the impact of the WTO on the number of case births is positive in the GATT/WTO data set, and that the models pass the likelihood ratio test. The Poisson and Negative Binomial models produce virtually identical results, though the AIC and SBC criteria are lower for the Negative Binomial. Explaining birth arrivals in terms of variables such as the existence of the WTO is harder than population, and this is reflected in the lower coefficients of
Table 3: Impact of WTO on Birth Rate of Trade Disputes

<table>
<thead>
<tr>
<th></th>
<th>GATT/WTO</th>
<th>USTR Section 301</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poisson</td>
<td>N-Bin.</td>
</tr>
<tr>
<td>WTO</td>
<td>1.6052</td>
<td>1.6052</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.229)***</td>
<td>(0.246)***</td>
</tr>
<tr>
<td>CONST</td>
<td>-1.6052</td>
<td>-1.6052</td>
</tr>
<tr>
<td>s.e</td>
<td>(0.173)***</td>
<td>(0.162)***</td>
</tr>
<tr>
<td>Goodness-of-fit test</td>
<td>288.48</td>
<td></td>
</tr>
<tr>
<td>GOF &gt; $\chi^2$</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Adjusted-R²</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-219.93</td>
<td>-210.05</td>
</tr>
<tr>
<td>LR (6)</td>
<td>65.41</td>
<td>45.12</td>
</tr>
<tr>
<td>LR &gt; $\chi^2$ (6)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>AIC</td>
<td>1.51</td>
<td>1.44</td>
</tr>
<tr>
<td>SBC</td>
<td>1.53</td>
<td>1.47</td>
</tr>
<tr>
<td>Sample Size</td>
<td>294</td>
<td>294</td>
</tr>
</tbody>
</table>

***Significant at 1%. **Significant at 5%. *Significant at 10%.

determination. Nevertheless, the coefficient estimates are statistically significant. According to them, the WTO raised the probability of new cases per month nearly fivefold, from .2 to 1.23

In contrast, the WTO does not appear to significantly influence the birth rate of new USTR Section 301 disputes. This agrees with our finding for population in Section IV.C. The base probability of a birth is .42 per month, and the WTO affects this probability by only 2 basis points.

B Lifespans: A Survival Analysis

One of the ways that the WTO can reduce adjudication costs is through shorter case timetables. We therefore tested whether stricter WTO time limits resulted in shorter caselength. In general, the study of event durations is an ideal candidate for a technique known as survival analysis. Often applied in the medical field, a standard question might be, What would the survival rates be for a group of bacteria, before and after infected patients are vaccinated by a particular drug? Survival rates might differ by the type of patient as well as their treatment.24 We are asking a similar question here: The patient is a country. Bacteria are trade dispute cases. The length of time a case is negotiated and litigated is the “survival” time or lifespan of the bacteria, and the change in treatment is the introduction of WTO rules and services to client nations.

Survival analysis is closely connected to the use of hazard rates. In the medical context, hazard rates are found by monitoring the rate of bacteria deaths after the introduction of a new vaccine.

23I.e., the difference between $e^{-1.6052+1.6052WTO=1} = 1.000$ and $e^{-1.6052+1.6052WTO=0} = .2008$.

24A number of interesting duration studies have appeared in different economic fields. Keenan (1985) provides one of the leading essays in the area, introducing survival analysis to explain the duration of labor strikes. Surprisingly, we could not find any study about the duration of international trade disputes. Nevertheless, it seems clear that a quantitative analysis of the issue can be particularly worthwhile to understanding the mechanics of international trade disputes.
In our setting, they correspond to the rate at which cases are solved after the WTO advent. If the WTO is effective, we expect that the number of solved cases relative to the monthly caseload (hazard rate) will increase and the duration of the cases (survival) will decrease after the application of WTO rules.

Following Neumann (1997), we used a nonnegative random variable, $T$, to describe the length of time until a trade dispute is finished. The cumulative distribution and the probability density functions of $T$, respectively, are

$$F(t) = \text{Prob}(T < t), 0 < t < \infty$$  \hspace{1cm} (16)$$

$$f(t) = \lim_{dt \to 0} \frac{\text{Prob}(t \leq T \leq t + dt)}{dt} = \frac{\partial F(t)}{\partial t}. \hspace{1cm} (17)$$

One can interpret $F(t)$ as the probability that a trade dispute will last no longer than $t$ days. The conditional probability that a trade dispute will last $t$ days or longer, called the survivor function, is given by

$$S(t) = \text{Prob}(T \geq t) = 1 - F(t). \hspace{1cm} (18)$$

From the survivor function one can obtain the instantaneous rate of event termination at $T=t$, conditional upon survival to time $t$, as the hazard function, $h(t)$:

$$h(t) = \lim_{dt \to 0} \frac{\text{Prob}(t \leq T < t + dt|T \geq t)}{dt} = \frac{f(t)}{S(t)}. \hspace{1cm} (19)$$

According to Heckman & Borjas (1980), the shape of the hazard function provides a characterization of the underlying stochastic process. There is a positive duration dependence—i.e., higher probability of case termination as time goes by—if $\partial h(t)/\partial t > 0$. The survival process is said to be memoryless when this first derivative equals zero, and exhibits negative duration dependence otherwise.

Variables that affect macroeconomic conditions and the volume of trade across countries might help to explain the duration of trade conflicts. Variables that we felt might matter included the bargaining power of the opponent measured by its gross trade share (GTS), the US degree of openness to trade (USOPEN), and the US trade balance (USTB), all evaluated at the final month of the dispute. The opponents’ gross trade share was measured by exports to, plus imports from the opponent country, scaled by US GDP. US openness to trade was measured as US imports plus exports divided by GDP. All monetary variables were in nominal form and converted to millions of US dollars. The monthly value of US imports and exports, as well as its annual GDP were extracted from the International Monetary Fund (2000). Specific information about the annual traded volume of imports and exports between the US and each country were extracted from the US Bureau of Census (1999) and the Organization for Economic Development and Cooperation (1999). Annual data were converted to monthly data through the application of finite geometric series indexes.

Besides these variables, we also tested the significance of indicator variables identifying whether the US was the plaintiff, whether the dispute ended in year of generalized foreign currency crisis, during the period of the Uruguay Round negotiations, or during a month of case overload (i.e., a period when the number of open cases was greater than the 80th percentile of its own series). None of the variables appeared to have significant explanatory power, except for the bargaining power of the opponent, GTS, which increased trade dispute lifespan for higher values.
B.1 The Cox Hazard Function Approach

The Cox proportional hazard model estimates a hazard function without imposing a parametric specification for the baseline hazard function $h_0$. The Cox hazard function is given by

$$h(t_j) = h_0(t)e^{\beta x_j}$$  \hspace{1cm} (20)$$

where $x_j$ is the vector of covariates. While the Cox model does not require a description of the baseline hazard function—an advantage—its disadvantage is that it estimates the effect of a treatment only as a deviation from the baseline. In our case, we are primarily interested in knowing what the effect of being a WTO case is relative to other cases so this is not an issue.

Table 4 presents the estimates for the Cox Proportional hazard model. We also provided the results of the test of the proportional hazard assumption, based on Grambsch & Therneau (1994), which checks if the log hazard ratio function is stable over time. If it is not, a stratified Cox regression would be required.

Table 4: Impact of WTO on Trade Dispute Lifespans: Cox Proportional Hazard Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard Ratio</th>
<th>Grambsch-Therneau Rank Test</th>
<th>ρ</th>
<th>$\chi^2$-stat</th>
<th>d.f.</th>
<th>Prob $&gt;\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>USTR$_{GATT}$ s.e.</td>
<td>0.5185 (0.128)**</td>
<td></td>
<td>0.100</td>
<td>0.98</td>
<td>1</td>
<td>0.322</td>
</tr>
<tr>
<td>USTR$_{WTO}$ s.e.</td>
<td>0.6023 (0.144)**</td>
<td></td>
<td>0.178</td>
<td>1.93</td>
<td>1</td>
<td>0.165</td>
</tr>
<tr>
<td>GTS s.e.</td>
<td>0.0407 (0.050)**</td>
<td></td>
<td>-0.142</td>
<td>2.46</td>
<td>1</td>
<td>0.117</td>
</tr>
<tr>
<td>Global Test</td>
<td></td>
<td></td>
<td>4.88</td>
<td>3</td>
<td></td>
<td>0.181</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-439.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald $\chi^2 (3)$</td>
<td>18.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald $&gt;\chi^2 (3)$</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>7.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBC</td>
<td>7.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td>118</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| GATT/WTO Cases | | | | | | |
|-----------------|-----------------|-----------------|-----|-----------------|
| WTO s.e. | 1.253 (0.282) | | 0.127 | 1.24 | 1 | 0.266 |
| GTS s.e. | 0.1583 (0.203) | | 0.055 | 0.22 | 1 | 0.643 |
| Global Test | | | 1.24 | 2 | | 0.537 |
| Log-likelihood | -294.20 | | | | | |
| Wald $\chi^2 (2)$ | 4.78 | | | | | |
| Wald $>\chi^2 (2)$ | 0.09 | | | | | |
| AIC | 6.97 | | | | | |
| SBC | 7.03 | | | | | |
| Sample Size | 85 | | | | | |

**Significant at 1%.  **Significant at 5%.  *Significant at 10%.

---

From the results in Table 4 we can see that the GATT, WTO, and trade share explanatory variables are significantly different from zero at the 5% level for the panel of USTR Section 301 cases. Moreover, the Grambsch-Therneau test for each variable shows that none of them violates the proportional hazard hypothesis as required by the Cox approach. The global test statistic for the whole equation also meets this requirement. The table indicates that being a USTR\textsubscript{GATT} case reduces the hazard rate to 52% of the benchmark hazard (USTR\textsubscript{ONLY}), while being a USTR\textsubscript{WTO} reduces it to 60% of the original hazard. This supports the finding that USTR Section 301 cases that resorted to the GATT mechanism were longer than USTR Section 301 cases that resorted to the WTO mechanism, but that both sets of cases were longer than USTR cases that did not resort either.

For GATT/WTO cases there are only two types of cases (GATT or WTO) and so only one treatment variable (WTO) representing deviation from baseline. Both the WTO treatment and the OPPONENT’S TRADE SHARE have the sign and dimension expected from the theoretical model. Both regressors pass the test for proportional hazard rate through time, as does the whole model in the global test. Nevertheless, neither regressor is statistically significant at the 5% level so that we cannot state with assurance that WTO cases have a higher hazard rate that GATT cases as the point estimates show.

B.2 Parametric Modelling

The success of the Cox model encourages us to hope that a fully parametric model might provide a good fit to the data, allowing us to estimate a base hazard rate as well as deviations from it. In parametric cases, the hazard function behaves according to a parametric probability distribution. For the present paper we considered the following distribution families: Weibull, Exponential, Gompertz, Gamma, Log-Normal, and Log-Logistic.

Traditionally, parametric survival analysis is subdivided into two types: Accelerated Failure-Time (AFT) models and Proportional Hazard Rate (PHR) models. The equation of AFT models is

$$\ln(t_j) = x_j\beta + \varepsilon_j$$

where $t_j$ is the survival time, $x_j$ is a vector of covariates, and $\varepsilon_j$ is the error term, which is assumed to be parametrically distributed. Here we assume the distributions are Gamma, Log-Normal, or Log-Logistic. The equation of PHR models is

$$h(t_j) = h_0(t)e^{x_j\beta}$$

where $h(t_j)$ is the hazard rate function, $h_0(t)$ is the baseline function assumed by us to be a Weibull, Exponential, or Gompertz, and $e^{x_j\beta}$ is the relative risk exposure.

Table 5 shows the results for the parametric estimation. Among USTR cases, the benchmark category is USTR\textsubscript{ONLY}. Therefore, the explanatory variables should be compared with this category. Among GATT/WTO cases, the benchmark category is the set of GATT cases.

The Weibull, the Exponential, and the Gompertz columns estimate the PHR format, and consequently do not provide any intercept. The Gamma, Lognormal, and the Log-Logistic columns estimate the AFT format, and provide their respective intercepts. Theoretically, the model with lower AIC and SBC criteria provides the best approximation of the “true” model among the parametric choices presented here.\textsuperscript{26}

\textsuperscript{26}For survival parametric models, the AIC and SBC criteria include all model-specific ancillary parameters.
The Gamma model seems to be the best representation of USTR Section 301 cases. Besides low AIC and SBC criteria (almost the lowest among the specifications), it also provides significant regressors at the 5% level. The Gamma model passes the Wald test for joint significance as well. If we choose the Gamma model as our compass, the marginal contribution of the opponent’s gross trade share (GTS) is positive and statistically significant—the higher the bargaining power of the opponent, the longer the expected lifespan of a dispute. Also, the contributions of resorting either to GATT or WTO assistance in a trade dispute are statistically significant and have a positive impact on the duration of the case. This result may be explained because those USTR Section 301 cases that resort to the GATT or WTO dispute settlement body are, on average, more complicated cases and it is natural to think they would have longer lifespans compared to USTR_{ONLY} cases. Finally, the estimates imply that WTO assistance is more effective than GATT assistance, in the sense that it solves cases faster, because the WTO contribution to the lifespan of USTR Section 301 cases is smaller than GATT’s contribution.

For GATT/WTO cases, Weibull and Gompertz models perform best in terms of the Wald test of joint significance of regressors. The Weibull model has lower AIC and SBC criteria, however, and therefore appears to be the better of the two. The Weibull model is the only one that provides statistically significant covariates, good Log-likelihood, plus low AIC, and SBC criteria. Looking just at the Weibull results, we see that being a WTO case increases the proportional hazard rate nearly 57%. This implies that the lifespan of WTO cases are, on average, shorter. The bargaining power of the opponent (measured by gross trade share, GTS) also enters positively. The higher the bargaining power, the lower the hazard rate and the longer the case survival time.

These results are consistent with the theoretical model and with the results from the Cox proportional hazard estimates. In the preferred Gamma and Weibull models, the regressors are all statistically significant. The sole exception is the coefficient for gross trade share in the Weibull model, and it is significant at the 5.9 percent level. Based on the model above, we estimate that the average lifespan of a trade dispute is 512 days for USTR_{ONLY} cases, 803 days for USTR_{WTO} cases, and 892 days for USTR_{GATT} cases. In other words, USTR cases that resorted to the WTO are approximately 3 months shorter than cases that resorted to GATT, though both are longer than cases that the USTR settled itself. In the GATT/WTO panel, WTO cases have an average lifespan of 520 days compared to the 673 day lifespan of GATT cases. The difference between them is nearly 5 months in favor of the WTO.27

For USTR cases, the forecast is done using the coefficients of the Gamma model, at the sample average value of GTS (0.1321). For the GATT/WTO panel, we transformed the coefficients of the Weibull model to the AFT scale (WTO=-0.2565, GTS=1.1695, and Constant=6.3350), and forecast at the sample average of GTS(0.1506). The last step for both panels was to take the exponential value of the obtained results. This provided the average lifespan in days for each forecast.

27The AIC and SBC criteria can be considered as a good proxy for the mean-square error approach, because they provide a balance between bias and variance—large models (with too many explanatory variables) provide less bias but more variance, while small models do the reverse. AIC and SBC select the more balanced model, penalizing the use of too many parameters at the same time rewarding large log-likelihood estimates.
### Table 5: Impact of WTO on Trade Dispute Lifespans: Parametric Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Weibull</th>
<th>Expo</th>
<th>Gompertz</th>
<th>Gamma</th>
<th>LogNormal</th>
<th>LogLog</th>
<th>USTR Section 301 Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USTR\textsubscript{GATT}</td>
<td>0.5575</td>
<td>0.6014</td>
<td>0.5693</td>
<td>0.5548</td>
<td>0.6186</td>
<td>0.6500</td>
<td></td>
</tr>
<tr>
<td>s.e</td>
<td>(0.144)**</td>
<td>(0.127)**</td>
<td>(0.140)**</td>
<td>(0.220)**</td>
<td>(0.244)*****</td>
<td>(0.215)*****</td>
<td></td>
</tr>
<tr>
<td>USTR\textsubscript{WTO}</td>
<td>0.6580</td>
<td>0.6838</td>
<td>0.6655</td>
<td>0.4500</td>
<td>0.6217</td>
<td>0.5398</td>
<td></td>
</tr>
<tr>
<td>s.e</td>
<td>(0.196)</td>
<td>(0.161)</td>
<td>(0.184)</td>
<td>(0.230)**</td>
<td>(0.194)*****</td>
<td>(0.175)*****</td>
<td></td>
</tr>
<tr>
<td>GTS</td>
<td>0.0250</td>
<td>0.0564</td>
<td>0.0361</td>
<td>2.5227</td>
<td>1.7148</td>
<td>1.7467</td>
<td></td>
</tr>
<tr>
<td>s.e</td>
<td>(0.034)**</td>
<td>(0.057)**</td>
<td>(0.046)*****</td>
<td>(1.157)**</td>
<td>(1.026)*</td>
<td>(0.980)*</td>
<td></td>
</tr>
<tr>
<td>CONST</td>
<td>5.9057</td>
<td>5.7313</td>
<td>5.7637</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s.e</td>
<td>(0.134)**</td>
<td>(0.154)*****</td>
<td>(0.118)*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-158.78</td>
<td>-162.40</td>
<td>-161.57</td>
<td>-156.76</td>
<td>-160.83</td>
<td>-156.36</td>
<td></td>
</tr>
<tr>
<td>Wald\textsubscript{\chi(3)}</td>
<td>15.64</td>
<td>19.25</td>
<td>15.09</td>
<td>18.74</td>
<td>16.57</td>
<td>20.33</td>
<td></td>
</tr>
<tr>
<td>Wald\textsubscript{\chi(3)}</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>2.74</td>
<td>2.80</td>
<td>2.79</td>
<td>2.71</td>
<td>2.78</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>SBC</td>
<td>2.81</td>
<td>2.87</td>
<td>2.86</td>
<td>2.78</td>
<td>2.85</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td></td>
</tr>
</tbody>
</table>

| GATT/WTO Cases      |         |      |          |       |           |        |                        |
|---------------------|---------|------|----------|-------|-----------|--------|                        |
|                     |         |      |          |       |           |        |                        |
| WTO                 | 1.5698  | 1.2130 | 1.4543  | -0.1895| -0.2227   | -0.0567|                        |
| s.e                 | (0.311)**| (0.151) | (0.213)*****| (0.121) | (0.140) | (0.122)|                        |
| GTS                 | 0.1279  | 0.3601 | 0.1938  | 1.0071 | 0.9246    | 0.8738 |                        |
| s.e                 | (0.139)* | (0.224)* | (0.156)** | (0.613)* | (0.704) | (0.576)|                        |
| CONST               | 6.1965  | 6.0612 | 6.0472  |        |           |        |                        |
| s.e                 | (0.135)**| (0.135)*****| (0.114)*****|        |           |        |                        |
| Log-likelihood      | -81.36  | -100.06 | -93.91  | -78.93 | -82.46    | -77.99 |                        |
| Wald\textsubscript{\chi(2)} | 10.58   | 5.23   | 12.56   | 5.46   | 4.36      | 2.60   |                        |
| Wald\textsubscript{\chi(2)} | 0.01    | 0.07   | 0.00    | 0.07   | 0.11      | 0.27   |                        |
| AIC                 | 1.96    | 2.40   | 2.26    | 1.90   | 1.99      | 1.88   |                        |
| SBC                 | 2.02    | 2.46   | 2.31    | 1.96   | 2.04      | 1.94   |                        |
| Sample Size         | 85      | 85     | 85      | 85     | 85        | 85     |                        |

***Significant at 1%. **Significant at 5%. *Significant at 10%.

### B.3 A Further Test of Model Validity

This section estimates Nelson-Aalen cumulative hazard functions, a non-parametric procedure. The Nelson-Aalen cumulative hazard function is given by the following maximum-likelihood estimate,

$$
\hat{H}(t) = \sum_{j|t_j \leq t} \frac{d_j}{n_j}
$$

(23)

where \(n_j\) is the number of open cases at time \(t_j\) and \(d_j\) is the number of finished disputes at time \(t_j\). The plot of the empirical hazard function displays useful information about the termination probability of cases. The estimated functions for both data panels are shown in the figures below.

For USTR Section 301 cases, we considered USTR cases that resorted to GATT or WTO assistance. The estimated hazard rate for the WTO cases is significantly higher than the hazard rate for cases that resorted to GATT assistance. This means that GATT cases had a longer lifespan than equivalent cases under the WTO and that the contribution of the WTO, therefore, was to reduce the length of USTR Section 301 cases that resorted to international dispute settlement.
Figure 7: Comparison of the Hazard Rate of Section 301 Cases Under GATT vs WTO

Figure 8: Comparison of the Hazard Rate of GATT Cases vs WTO
mechanisms. This estimate and finding supports the Cox semi-parametric and the parametric results of the previous sections.

For GATT/WTO cases, we compared the hazard function of WTO cases against GATT cases. The resulting functions reveal an interesting pattern. If we divide the lifespan grid into three parts, the first from zero to 320 days, the second 321 to 700 days, and the third from 701 to 3600 days, we observe three different relationships. In the first and third intervals, the WTO hazard rate is above the GATT hazard rate. In the second interval, the result is reversed. This double-crossing property may partly explain why we had difficulty affirming that WTO cases are statistically significantly shorter than GATT cases in the whole sample in the semi-parametric Cox hazard function estimation. The death hazard for young cases is greater if they are under WTO-auspices, as is the death hazard for the longer and presumably more contentious disputes that are aged more than 700 days. In the middle, the death hazard is not greater for WTO cases, if only because cases that would have become middle-aged under other treatments have already died before their 321st day.\footnote{A suggested parametric approach to deal with such alternation along the sample is the use of quantile regressions. Survival analysis quantile regression is being currently studied by Koenker and Bilias (2001).}

VII. Conclusion and Summary

We constructed a theoretical model of international trade disputes and compared its main implications to the results of an empirical analysis of international trade disputes from 1975 to 2000 involving the US as complainant or defendant. We modeled trade disputes as the rational outcome of a finite extensive form noncooperative repeated game with one-sided incomplete information. The game-theoretic model involved a complainant ($\alpha$), the infringer ($\beta$), and the random arrival of opportunities to infringe. The infringer’s choice to infringe was based on economic self interest and the infringer’s beliefs about how the complainant would respond. Once an infringement has been discovered, $\alpha$’s trade representative decides whether to accept the case for processing (prosecution), and chooses $\alpha$’s strategy regarding a subsequent negotiated settlement with the infringer and whether to proceed to WTO adjudication if settlement cannot be reached. The complainant chose to pursue litigation of cases out of economic self interest, taking into account the value of maintaining a reputation as a country that defended its trade rights. A credibility condition and an infringement constraint were identified as the crucial elements for the game. Without a credible reputation for prosecuting cases, $\alpha$ would allow other countries to feel free to infringe its WTO rights with impunity. If so, negotiated settlements would never be reached because the defendant would have no incentive to accept them. The game was solved for the equilibrium strategies of infringer and complainant in terms of the costs of litigation, and the costs, benefits, and characteristics of the infringement and WTO adjudication process. The model was examined for its implications for changes to several types of costs to players. The main finding was that lower adjudication costs imply more decisions to infringe and more dispute litigations in game equilibrium. Since the WTO rules impose strict timetables on trade disputes, the model suggests that the activity of the WTO should result in more cases of shorter lifespan.

Table 6 summarizes the main results of an econometric evaluation of the population, birth, and lifespan of trade disputes of the last 25 years. Three types of data were examined: USTR Section 301 disputes (1975-2000), GATT trade disputes (1975-1994) involving the US as complainant or defendant, and WTO trade disputes (1995-2000) involving the US as complainant or defendant. We began by fitting a first order Poisson Markov regression to the data on population of trade disputes.
Our estimates indicated that the WTO increased the population of GATT/WTO cases between 25% and 33% depending on which model one used. In contrast to the statistical significance of the WTO in the first data set, the WTO did not appear to increase the population of USTR Section 301 disputes, a puzzle that we suggested may partly be explained by the discovery that after the WTO’s inception there was a statistically significant increase in the proportion of USTR cases that were treated under WTO auspices compared to the proportion treated under GATT.

The conclusion of the birth experiments was that the WTO does appear to have increased the number of cases—the expected number of new cases rose from .2 to 1 per month. This is consistent with the view that the WTO lowered adjudication costs and shortened caselength. An increase in case births also agrees with the predictions of the simple repeated game model.

We also implemented a survival analysis to study the effect of the WTO on case lifespans, starting by estimating a Cox proportional hazard model. The Cox model passed a validity test based on the methodology suggested by Grambsch & Therneau (1994). The results for USTR Section 301 trade disputes indicated that USTR cases that resorted to the WTO mechanism were shorter than those that resorted to GATT, but that both sets of cases were longer than USTR cases that did not resort to either. Point estimates for the sample of GATT/WTO cases also implied that WTO cases had a higher hazard rate (shorter lifespan), but the coefficient was not statistically significant. The Cox model was followed by estimating six parametric hazard models based on Weibull, Exponential, Gompertz, Gamma, Log-Normal, and Log-Logistic distributions. A similar finding emerged that the lifespan of WTO cases are significantly shorter, on average, than their GATT counterparts. Cases handled by the WTO were 3 months shorter in the USTR data set and 5 months shorter in the GATT/WTO data set. The estimates showed that increased
bargaining power of the trade dispute opponent (here measured by the size of its trade relative to US GDP) increased the dispute lifespan.

Finally, we calculated Nelson-Aalen cumulative hazard functions. A comparison of hazard rates showed that the WTO reduced the lifespan of USTR Section 301 cases that resorted to international dispute settlement mechanisms, consistent with the Cox semi-parametric and parametric results. We found that the hazard functions displayed a double-crossing property, indicating that the WTO performed better at terminating cases under 320 days and over 700, but that in the middle, it was less effective.

In summary, it appears that the data support the view that the WTO has shortened the lifespan of typical trade disputes, but increased their birthrate. The US is using the WTO more heavily than it used GATT, and trade disputes with larger opponents, measured by the size of their trade relative to US GDP, last longer. The theoretical model and the econometric models presented here, therefore, appear to provide a promising beginning to interpreting the impact of WTO on trade disputes. Ultimately, one hopes that further improvements in our understanding may encourage refinements and improvements in the dispute mechanism itself.

Appendix

- **Proportional Hazard Cases**

  The Weibull and the Exponential proportional hazard functions are given by equation 22, where \( h_0(t) = pt^{p-1} \) for the Weibull case and \( h_0(t) = 1 \) for the Exponential case. \( p \) is a shape parameter estimated from the data. The Gompertz proportional hazard model is given by

  \[
  h(t_j) = e^{\gamma t_j + x_j \beta}
  \]

  where \( \gamma \) is a parameter estimated from the data.

- **Accelerated Failure Time Cases**

  For the Gamma AFT model, the survivor function is given by

  \[
  S(t) = 1 - I(\kappa, \kappa \cdot exp\left(\frac{z}{\sqrt{\kappa}}\right))
  \]

  where \( z = \frac{\ln(t) - \lambda}{\sigma} \), and \( \kappa \) and \( \sigma \) are ancillary parameters estimated from the data. The Gamma hazard function is quite flexible. For example, when \( \kappa = 1 \), the Gamma hazard has the Weibull shape; when \( \kappa = 1 \) and \( \sigma = 0 \), it has the Exponential shape, and when \( \kappa = 0 \) it has the Lognormal shape.

  In the case of the Lognormal AFT model, the survivor and hazard functions are, respectively,

  \[
  S(t) = 1 - \Phi\left(\frac{\ln(t) - \mu}{\sigma}\right)
  \]

  and

  \[
  h(t) = \frac{1}{t\sigma \sqrt{2\pi}} exp\left(-\frac{1}{2\sigma^2} (\ln(t) - XB)^2\right)
  \]

  where \( \sigma \) is the standard deviation, \( \Phi(\cdot) \) is the standard normal cumulative distribution, and \( X \) are the covariates.
The Log-Logistic survivor and hazard functions are

\[
S(t) = \frac{1}{1 + (\lambda t)^\gamma}
\]

(28)

and

\[
h(t) = \frac{\lambda t^{\gamma - 1}}{\gamma (1 + (\lambda t)^\gamma)}
\]

(29)

where \( \lambda = e^{-x_j \beta} \) and \( \gamma \) is estimated from the data.

References


[23] Koenker, Roger, and Bilias, Yannis. (2001) Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments, mimeo, University of Illinois at Urbana-Champaign, Department of Economics.


