# Gravitation or discrimination? Determinants of litigation in the World Trade Organisation

THOMAS SATTLER[1] & THOMAS BERNAUER[2]
[1]*School of Politics and International Relations, University College Dublin, Ireland;* [2]*Centre for Comparative and International Studies, ETH Zürich, Switzerland*

**Abstract.** The strong presence of large countries in World Trade Organisation (WTO) dispute settlement and the absence of very poor ones have raised concerns that increasing legalisation in the global trading system has not diminished discrimination against less powerful countries as much as expected. This article examines dispute initiations in all WTO member state dyads in 1995–2003 to shed more light on this issue. The analysis suggests that the main driver of dispute initiation is a gravitational one: larger economies and bigger traders are more likely to become involved in trade disputes primarily because their economies are more diversified, and also because greater market size makes them more attractive targets of litigation. While evidence is not found for discriminatory effects against countries with small legal capacity, the results of the article point to a more complex form of power bias – namely a preponderance effect. They suggest that disputes among country dyads including a much more powerful defendant than complainant or vice versa are dealt with outside the WTO. This finding is potentially worrying because it is, arguably, easier to reduce legal capacity differences than to reduce power differences.

One of the most important issues in contemporary international relations research is whether and under what conditions the rule of law that is characteristic of modern democracies can be extended beyond its cradle, the nation-state, to solve international problems through legal means (Goldstein et al. 2000: Abbott et al. 2000; Zangl 2008). The dispute settlement mechanism of the World Trade Organisation (WTO) offers a great opportunity to study this issue. This mechanism is unique in that it offers the international community a tightly structured venue for resolving trade disputes through a judicial system rather than conventional state-to-state bargaining in which more powerful countries are usually better positioned for obtaining their preferred outcome.

One of the key issues in trying to understand the extent to which legalisation in the global trading system has progressed, or conversely remains incomplete, concerns power and capacity biases in WTO dispute settlement. Descriptive evidence on WTO disputes shows that the major economic powers such as the United States, the European Union (EU) and Japan are frequent users of this mechanism, while the least developed countries are almost absent.

Does this indicate an important power and capacity bias in WTO dispute settlement, in the sense that this system discriminates against smaller and poorer countries – that is, countries that are less powerful and have smaller capacities to litigate?

We contribute to the existing literature on the subject (Allee 2004, 2005; Bernauer & Sattler 2006; Busch et al. 2008; Davis & Blodgett Bermeo 2009; Guzman & Simmons 2005; Horn et al. 1999; Kim, 2008) in three novel ways. First, we estimate and explicitly compare the relative importance of discriminatory and gravitational effects. Three distinct explanations emphasising discrimination argue that variation in countries' involvement in WTO disputes is primarily due to differences in power and legal capacity. The fourth explanation, the gravity argument, holds that larger economies are more likely to become involved in trade disputes primarily because their economies are more diversified and greater market size makes them more attractive targets of litigation. Second, besides analysing more conventional, direct power effects, we also study a more complex, indirect power effect not considered in other studies on trade disputes: power preponderance. Third, we use a combination of standard and zero-inflated count models. The latter allow us to distinguish between dispute-enabling and dispute-promoting factors.

The empirical analysis relies on a dataset that includes all directed WTO member state dyads from 1995 to 2003. The results strongly support the gravity argument – that is, we find that large economies are pushed toward WTO disputes, both as complainants and defendants, more often than the larger trade volume of large countries would suggest. In contrast to other studies, our results do not support the legal capacity hypothesis. However, we identify a particular type of power-related bias that has not been analysed before. Dyads with larger differences in power between complainant and defendant – in either direction – are less likely to become involved in a WTO dispute. This power preponderance effect suggests that more powerful countries may obtain concessions from less powerful countries outside the WTO, and/or that less powerful countries may abstain from formal WTO litigation for fear of reprisals.

We conclude from these findings that optimism about legalisation in the global trading system is warranted only in part (e.g., Zangl 2008). The positive news is that gravity effects, which from a legalisation perspective are rather unproblematic, are strong, and that our results do not support concerns about discriminatory legal capacity bias. The negative news is that we observe a considerable preponderance effect. This effect suggests that disputes among country dyads characterised by strong power asymmetry are dealt with outside the WTO. This finding is potentially worrying because it is much easier to reduce differences in legal capacity (e.g., by providing legal support for poor

countries through international organisations and nongovernmental organisations) than to correct for power differences.

## The issue

Even though the WTO dispute settlement system has been heralded as an institutional innovation that puts right before might, critics maintain that some countries in this system are 'less equal' than others. In particular, less powerful countries may not initiate legal action against a more powerful country because they fear costly retaliation. Discriminatory effects may also emanate from the fact that involvement in a WTO case can be very costly because it requires specialised legal staff and disputes may last several years (Bown 2005a; Davis & Blodgett Bermeo 2009; Horn et al. 1999; Kim 2008; Busch et al. 2008; Smith, 2004).

Is WTO dispute settlement biased against less powerful and poorer countries? A first, descriptive look at the available data on WTO dispute initiations produces an ambiguous picture. As shown in Figure 1, industrialised countries are by far the heaviest users of the dispute settlement system. They initiated more than 60 per cent of the WTO disputes and acted as defendants in more than 70 per cent of the cases. In contrast, less than 1 per cent of the disputes were initiated by least developed countries, and least developed countries never appear as defendants. Developing countries as a whole have made consistent use of the system, however.

The picture changes if one weights participation in dispute settlement by shares in world trade. In 2000–2006, merchandise exports by developing
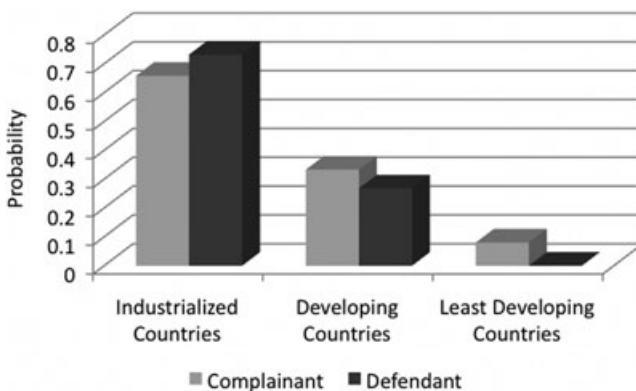


*Figure 1.* Dispute initiators and targets in the WTO system, 1995–2006.
Source: Horn & Mavroidis (2008).

countries were around 30–35 per cent of total global trade (WTO 2006). This share corresponds, by and large, to developing countries' share of dispute initiations in the WTO. These empirical patterns do not offer a conclusive answer concerning the extent of power- and capacity-related biases in WTO dispute settlement. Such an answer requires more nuanced theoretical arguments and systematic empirical testing.

## Discrimination or gravitation?

Relative power and legal capacity play a prominent role in studies that account for litigation in the global trading system and variation of participation in dispute settlement (e.g., Conybeare 1985; Guzman & Simmons 2005; Davis & Blodgett Bermeo 2009; Busch et al. 2008; Kim, 2008). Based on this literature, we distinguish three arguments that reflect potential discriminatory effects in the WTO: a legal capacity argument, a standard power argument and a power preponderance argument. The latter has not been analysed in the trade disputes literature before. By combining the literatures on peace research and trade conflicts (Horn et al. 1999; Hegre 2008), we then add a gravitational argument. This argument contends that larger economies are more likely to become involved in trade disputes for reasons unrelated to discrimination.

This section identifies and compares alternative mechanisms related to country size and capacity that may generate the empirical patterns just discussed. It lays the ground for different empirical specifications that all include economic size and legal capacity (subsequent empirical section). We define the two arguments based on *relative* size as power arguments because the crucial component of the mechanism that leads to conflicts involves coercion by one country against the other. This is not the case for the gravity argument, which refers to *absolute* country size. The gravity mechanism that leads to trade conflicts is based on one-sided cost-benefit calculations and does not involve pressure from one side on the other.

### *Discrimination associated with differences in power and legal capacity*

Studies on international trade disputes commonly assume that the distribution of power within a country dyad affects the trade policy of one country toward the other. In particular, the more powerful country is presumably better positioned to impose its will on the less powerful one. For example, in his analysis of French-Italian trade disputes in the 1880s and 1890s, Conybeare (1985: 147) notes that 'big powers can coerce small powers'. Guzman and Simmons (2005:

557) state that 'politically weak countries will refrain from filing complaints against politically powerful states for fear of costly retaliation'.

The logic of this standard power argument is that if the complainant country is more powerful than the target country, it can impose greater costs on the target with less harm to its own economy. It thus faces lower costs of retaliation by the defendant when it considers whether or not to initiate a dispute against a less powerful country. Moreover, countries that obtain a favourable WTO verdict are themselves responsible for enforcement; more powerful complainants also have more power to enforce verdicts in their favour, while more powerful defendants are better able to resist the implementation of verdicts against them. A complainant is therefore more likely to get what it wants if the target country is less powerful. Vice versa, countries are more likely to refrain from initiating a WTO dispute against a more powerful country, even if they have reasons to believe that they could win the legal case.

A related version of this mechanism, which we call the 'power preponderance argument' and which has so far not been examined in this context, takes into account that states are likely to interact outside the WTO before they take legal action inside the organisation's dispute settlement system. In that case, countries of different power will first negotiate bilaterally about barriers to trade, and powerful states, whether complainants or defendants, will try to obtain concessions from less powerful states. More powerful complainant countries will ask weaker defendant countries to lift disputed trade restrictions without WTO proceedings. Similarly, more powerful defendant countries will resist demands by weaker complainant countries to lift trade restrictions, but weaker complainants may not initiate a formal WTO dispute because they fear reprisals. The probability of an 'out-of-court' settlement then increases with larger power asymmetry – in either direction – between the complainant and defendant, and WTO dispute initiation becomes less likely. This hypothesis resembles the power preponderance hypothesis in research on armed interstate conflict.

The second discriminatory factor besides power is captured by the legal capacity argument. Greater legalisation has increased the complexity of the WTO dispute settlement process, and many observers of the WTO suspect that countries with smaller legal capacity are less likely to be able to use improved opportunities for taking legal action in the WTO (Horn et al. 1999; Kim 2008; Busch et al. 2008). Poor countries may not have the financial means and human resources to prepare and follow through with a WTO dispute. Taking legal action in the WTO can be very costly in several respects. It requires that countries monitor other countries' trade policies and import restrictions, collect evidence on policies and practices that violate international trade law, and hire legal staff, sometimes for years. Countries with larger legal capacity

may also be more likely to resist attempts to settle a dispute outside the WTO and accept a greater risk that the complainant will take the dispute to there. Thus, we should expect countries with a larger legal capacity to be more active users of the dispute settlement mechanism.

Some authors have pointed to circumstances under which the discriminatory effect of low legal capacity may diminish. Previous experience with WTO dispute settlement can equip poor countries with expertise for future cases (Davis & Blodgett Bermeo 2009). This argument implicitly suggests that, unlike power relations in a dyad, legal capacity is not strictly exogenous to a country's propensity to initiate or become a target of legal action in the WTO. For instance, poor countries facing trade restrictions imposed by an important export destination may conclude that the expected benefits from taking legal action and eventually gaining better access to that market will be greater than the costs of acquiring a larger legal capacity. Once they have made this investment, the barrier to participation in future disputes is lower.

*Larger economies gravitate into trade disputes*

An alternative explanation of the empirical patterns outlined above contends that most variation in dispute initiation may not be due to discrimination, but rather is caused by gravitation. The gravitational argument holds that dispute propensity is caused primarily by economic size and trade volume. It postulates, moreover, that large economies participate in disputes disproportionately often – that is, more often than their larger trade volume would suggest – because their economies are more diversified and because greater market size makes them more attractive targets of litigation. Although the gravity argument relies on country size, we treat it separately from the two power arguments above. The reason is that the power logic, unlike the gravitation logic, assumes coercion to be the driving force. In contrast, the gravity argument assumes a one-sided cost-benefit calculation by each country that does not involve pressure by one country on the other.

The idea that large countries gravitate towards international conflicts has its roots in the gravity model of trade. This model explains bilateral trade flows as a function of the economic mass and distance between two countries (Bergstrand 1985). Translating this model to international armed conflict has generated the proposition that larger countries (in absolute terms) are more likely to become involved in militarised disputes (Hegre 2008). This 'gravitation' effect in interstate conflict implies that large (small) countries are involved in more (less) conflicts than the distance between two potential opponents suggests.

The gravity logic applies to trade disputes as well (see also Horn et al. 1999). Larger economies have more opportunities to become involved in a trade conflict because they tend to be economically more diversified than smaller economies and usually have a larger range of industries. Large economies (e.g., the United States, Japan or Germany) produce and trade all major goods and services, from agricultural products to steel and ships to sophisticated financial services and consumer electronics. Even highly developed small economies (e.g., Switzerland or New Zealand) do not have some of the industries that still play a considerable role in large economies, such as heavy industries. Consequently, it is more likely that a large exporting country will encounter disputable trade restrictions in some other country. Conversely, given the greater diversity of imports it is also more likely that a trade partner will be negatively affected by an import restricting measure of the large economy and decide to challenge this measure in the WTO.[1]

The economic size of countries also affects motivations to become involved in a trade dispute. Large economies are particularly attractive targets for potential complainants because of their large market size, which ultimately affects the economic stakes of a case (Guzman & Simmons 2005; Allee 2004). If the WTO issues a verdict in favour of the complainant and the defendant complies, exporters from the complainant country can gain more when the destination market is larger. The incentive to take legal action is therefore particularly strong when suspected violations of trade law occur in a large economy.

The gravity logic postulates that both bilateral trade and economic size promote dispute initiation, and that large economies experience more disputes than predicted by their larger trade volume alone. Relating the gravity model of disputes to the classic gravity model, the role of trade in the dispute model corresponds to the role of distance in the classic model while economic mass is the same in both models.[2]

Table 1 summarises the predictions and highlights the differences among the four explanations of trade disputes. For instance, both the power preponderance and the gravity mechanisms suggest that the conflict propensity between two large countries should be high. Yet unlike the preponderance argument, which implies that the probability of conflict should be large for two small countries, the gravity logic suggests that the propensity for this dyad should be low. The standard power argument predicts an intermediate conflict propensity for the small countries dyad.

Distinguishing these four types of effects is important not only from an analytical perspective, but also from a policy perspective. If involvement in WTO trade disputes is driven primarily by gravitation, this is good news for proponents of legalisation in the global trading system. We could then con-

*Table 1.* Summary of predictions by alternative explanations

| Dyad characteristics | Predicted conflict propensity | | | |
|---|---|---|---|---|
| Complainant/ defendant | Standard power | Power preponderance | Legal capacity | Gravity |
| Small/Small | Intermediate | High | Low | Low |
| Small/Large | Low | Low | Low | Intermediate |
| Large/Small | High | Low | High | Intermediate |
| Large/Large | Intermediate | High | High | High |

Notes: Dyad characteristics in the standard power, power preponderance and gravity models refer to countries' economic size. In the legal capacity model, they refer to a country's ability to mobilise legal resources to get involved in a WTO dispute.

clude that the WTO process is effective in putting right before might. To the extent that we observe discriminatory effects emanating from differences in power and legal capacity, the news will be less negative if such effects emanate from differences in capacity rather than power since it is much easier to design and implement policies that help reduce differences in legal capacity as opposed to differences in power.

## Empirical design

### Approach

A comprehensive study of dispute initiation and its determinants must pay attention to all potential conflict dyads (country pairs), including those that have not experienced a dispute. Moreover, testing arguments concerning power requires that we use explanatory variables that capture the characteristics of both the complainant and the defendant country. While the complainant ultimately decides whether or not to initiate a WTO dispute, this decision also depends on the defendant country's willingness to comply with demands by the complainant.

The design of our analysis thus follows the approach that has been used to study gravity and power effects in interstate armed conflict. The dataset includes all *directed* WTO member state dyads from 1995 to 2003. Each dyad appears twice because the dependent variable measures not only the number of trade disputes in a dyad, but also which country is the complainant and which is the defendant. We restrict the analysis to the 1995–2003 period because of missing data for explanatory variables after 2003. We then drop

'irrelevant dyads'; in our context, these are dyads that have zero bilateral trade in either direction.[3]

Our empirical approach is based on the implicit assumption that the underlying conflict potential – that is, the probability of encountering a disputable measure in a partner country – increases with greater trade.[4] We believe that this assumption is reasonable from a theoretical perspective, but also empirically examine the relationship between conflict potential and trade.[5] The estimated relationship between trade and conflict potential not only is very strong, but also very precise. This indicates that very few countries deviate from this rule. We conclude from these results that the assumption about the positive and systematic relationship between trade and conflict potential underlying our analysis is reasonable.

We implement the empirical analysis in two steps. The first relies on a cross-sectional design that examines the number of trade dispute initiations a directed dyad experienced in the 1995–2003 period. We opt for the cross-sectional approach because dispute initiations among WTO dyads are very low probability events – even after excluding irrelevant dyads. The cross-sectional analysis reduces to some extent the large number of 'no events' in the dataset. In the second step we re-examine the cross-sectional results with a dataset that includes dispute initiations across dyads and over time.[6]

*Variables*

The dependent variable measures how many trade disputes a WTO member country initiated against another member country in a given year. This definition requires that we analyse annual country pairs and split disputes filed under the WTO dispute settlement mechanism by more than one country into dyads. This approach follows a common practice in the existing literature (e.g., Busch 2000). The reason is that disputes initiated by several countries can be settled (or escalated) bilaterally. The EU is treated as a single actor because its members generally pursue a common trade policy in the WTO context. A dispute initiation is coded as such if a formal request for consultations under the WTO dispute settlement system was made.[7]

The key explanatory variables in the gravity model are the log of bilateral trade and the log of economic size of potential complainants (country A) and defendants (country B). Trade is the sum of imports and exports between two countries in billions US\$. Economic size is the GDP of the respective country. The data are taken from Gleditsch (2002) and updated with data from the International Monetary Fund Directions of Trade Statistics and the Penn World Tables. We expect the coefficients on trade and economic size to be positive, indicating that greater traders and larger countries are involved in

more disputes. We also examine whether the causal chain connecting complainant's economic size to dispute propensity works through export diversity as postulated by the gravity mechanism. As a measure of export diversity, we use the number of product categories in which a country exports (according to the four-digit SITC classification). This indicator is constructed using data from Feenstra et al. (2005).

The key explanatory variable in the conventional power model is economic size in *relative* terms. 'Relative power' is the difference between the log GDP of the complainant and log GDP of the defendant. It takes the value 0 if both countries are equally powerful. Negative values show that the defendant is more powerful than the complainant; positive values indicate that the complainant is more powerful. To test the preponderance hypothesis, we measure 'power asymmetry' in terms of the absolute value of the relative power variable. High values indicate strong power asymmetry, but the variable does not specify which of the two countries is more powerful. We thus follow the standard procedure in international conflict research to measure power differences among countries (see, e.g., Hegre 2008). Although all variables are based on measures of GDP, the dyadic nature of the two power variables reflects the country's power in relation to a specific opponent, which is fundamentally different from absolute GDP. The coefficient on relative power should be positive, indicating that large countries initiate disputes against small ones more often than vice versa. The coefficient on power asymmetry should be negative, indicating that with greater asymmetry, the probability of a dispute decreases.

We also examine whether 'relative income' and 'income asymmetry' have an effect on dispute initiation because differences in income may proxy for differences in power as well. These two terms are defined in the same manner as the relative power variable, using the log of GDP per capita of the two countries. We include bilateral trade in all power models as a control variable.

The key variable in the capacity model is measured in terms of income and WTO delegation size – that is, the log of GDP per capita [Log(Inc. A), Log(Inc. B)] and the number of delegates to the WTO in Geneva (Delegates A, Delegates B). Both indicators have been used as proxies for legal capacity in previous studies. GDP per capita reflects the idea that poor countries have less means to engage in trade disputes (Bown 2005b). Delegation size captures the means invested in legal capacity more directly.[8]

The control variables are 'democracy of A and B', measured with the combined scores from Polity IV for each county in a dyad (Marshall et al. 2002), and 'retaliation' in models using cross-section time-series data. In line with other work, we expect that democratic institutions have a positive effect

on dispute initiation (Allee 2004; Rosendorff 2005; Busch 2000; Davis & Blodgett Bermeo 2009; Reinhardt 1999). The retaliation variable, which is expected to have a positive effect, measures whether country A was the target of a dispute initiation by B before it initiated a dispute against that country (Allee 2005). Its value is 1 if the defendant country (B) initiated a dispute against the complainant (A) during the same or the previous year, and 0 otherwise.[9] Table 2 presents the descriptive statistics of the variables.

*Method*

We combine previous empirical modeling approaches by focusing on the characteristics of both complainant and defendant countries (e.g., Reinhardt 1999) and accounting for the exact number of disputes that were initiated (Davis & Blodgett Bermeo 2009). Moreover, we use a combination of standard and zero-inflated count models. The latter help in dealing with excess zeros in the data and allow us to distinguish dispute-enabling and dispute-promoting factors.

We start with negative binomial models to estimate the expected number of trade dispute initiations by one country against another (King 1988). The data

*Table 2.* Summary statistics

| Variable | Observations | Mean | Standard deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Log(Trade) | 64,073 | −4.90 | 3.18 | −30.26 | 6.09 |
| Log(GDP A) | 64,073 | −2.69 | 1.90 | −7.42 | 2.32 |
| Log(GDP B) | 64,073 | −2.69 | 1.90 | −7.42 | 2.32 |
| Relative Power | 64,073 | 0.00 | 2.88 | −9.61 | 9.61 |
| Relative Income | 64,073 | 0.00 | 1.61 | −4.62 | 4.62 |
| Power Asymmetry | 64,073 | 2.30 | 1.72 | 0.00 | 9.61 |
| Income Asymmetry | 64,073 | 1.30 | 0.94 | 0.00 | 4.62 |
| Log(Income A) | 64,073 | 1.66 | 1.12 | −1.08 | 3.59 |
| Log(Income B) | 64,073 | 1.66 | 1.12 | −1.08 | 3.59 |
| Delegates A | 62,289 | 5.97 | 4.53 | 0.00 | 23.00 |
| Delegates B | 62,305 | 5.98 | 4.53 | 0.00 | 23.00 |
| Democracy A | 64,073 | 4.33 | 6.09 | −10.00 | 10.00 |
| Democracy B | 64,073 | 4.33 | 6.07 | −10.00 | 10.00 |
| Retaliation | 64,073 | 0.00 | 0.05 | 0.00 | 1.00 |

Note: The table shows the summary statistics for the directed annual dyads (i.e., for the time-series cross-section dataset).

structure in this first step is cross-sectional – that is, we use 1995–2003 averages for the explanatory variables and the cumulated number of disputes in a dyad in that time period.[10] The use of a cross-sectional approach alleviates the problem of excess zeros – that is, the large number of non-events in the time-series cross-section dataset. Hence it mitigates the difficulties in properly estimating conflict propensity. The number of conflicts per dyad during the pre-specified period captures the overall conflict propensity of a dyad. According to the arguments outlined above, the expected number of conflicts for the pre-specified period should vary with the level of GDP, the power distribution in a dyad and so on. These explanatory variables vary only very little over time, which means that the averages represent the characteristics of a dyad for the period of analysis well.[11] The time-series cross-section analysis described in the next paragraph assesses the robustness of the cross-sectional results. The risk of becoming involved in a WTO dispute is higher the longer a country is a WTO member. To account for differences in exposure time, our models include the log of the duration (years) of the complainant's WTO membership. The coefficient for this variable is fixed at 1.

To address the excess-zeros problem in the time-series cross-section analyses, we employ zero-inflated negative binomial models (Greene 1994). The first stage of zero-inflated models (the inflation equation) uses a binary specification to estimate whether a dispute is possible (i.e., whether there is a positive probability of a dispute). The second stage (the conflict equation) accounts for variation in the number of disputes initiated among those dyads that have a positive probability of a dispute. This approach has several advantages. The zero-inflated models split the dispute initiation process into two stages for statistical and theoretical reasons. It is unrealistic to assume that all dyads have a strictly positive probability of a trade dispute, as the standard models assume. Countries that trade very little (even if more than nothing) are very unlikely to become involved in a dispute over trade restrictions. Since it is not obvious *ex ante* at what trade level a dyad becomes 'relevant', the two-step procedure separates observations with zero probability of dispute initiation from those with a positive probability.[12]

## Results

Table 3 shows the results of the cross-sectional analysis. The gravity model in the first column offers strong support for the argument that large economies gravitate towards disputes. Absolute economic size of both the complainant and the defendant has a strong effect on the conflict propensity of a dyad. Large economies initiate more disputes than small ones, and they are more

Table 3. Cross-sectional (negative binomial) models of gravity, power and capacity

| | Gravity model | Standard power | Power preponderance | Capacity model I | Capacity model II | Full model I | Full model II |
|---|---|---|---|---|---|---|---|
| Log(Trade) | 0.424*** (0.086) | 0.700*** (0.038) | 0.707*** (0.037) | 0.786*** (0.044) | 0.704*** (0.052) | 0.454*** (0.071) | 0.521*** (0.083) |
| Log(GDP A) | 0.277** (0.118) | | | | | 0.312*** (0.110) | 0.234** (0.097) |
| Log(GDP B) | 0.367*** (0.120) | | | | | 0.422*** (0.110) | 0.343*** (0.093) |
| Rel. Power | | −0.034 (0.032) | | | | | |
| Rel. Income | | −0.068 (0.066) | | | | | |
| Power Asym. | | | −0.238*** (0.069) | | | −0.231*** (0.063) | −0.243*** (0.063) |
| Income Asym. | | | 0.100 (0.126) | | | | |
| Log(Inc. A) | | | | −0.470*** (0.114) | | | −0.340*** (0.131) |
| Log(Inc. B) | | | | −0.240** (0.096) | | | −0.207* (0.109) |
| Delegates A | | | | | −0.016 (0.015) | −0.028 (0.021) | |
| Delegates B | | | | | 0.013 (0.018) | −0.021 (0.022) | |
| Democracy A | 0.152*** (0.026) | 0.168*** (0.030) | 0.154*** (0.024) | 0.202*** (0.037) | 0.156*** (0.029) | 0.158*** (0.025) | 0.195*** (0.035) |
| Democracy B | 0.119*** (0.023) | 0.122*** (0.024) | 0.146*** (0.026) | 0.155*** (0.028) | 0.120*** (0.026) | 0.119*** (0.024) | 0.140*** (0.028) |
| Constant | −6.117*** (0.345) | −6.747*** (0.322) | −6.370*** (0.387) | −5.631*** (0.441) | −6.593*** (0.427) | −5.032*** (0.525) | −4.809*** (0.443) |
| $\hat{\alpha}$ | 1.65 | 2.58 | 2.18 | 2.48 | 2.60 | 1.38 | 1.57 |
| N | 9,928 | 9,928 | 9,928 | 9,928 | 9,123 | 9,111 | 9,928 |
| $\chi^2$ | 658.53 | 494.83 | 512.53 | 507.75 | 486.96 | 690.05 | 809.21 |
| Prob > $\chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| BIC | 1,184.12 | 1,203.64 | 1,189.07 | 1,192.92 | 1,201.15 | 1,185.97 | 1,183.11 |
| Log likelihood | −559.85 | −569.61 | −562.33 | −564.25 | −568.66 | −547.40 | −545.54 |

Notes: Robust standard errors cluster on dyad and are listed in parentheses. All models include the log of duration to account for differences in exposure time.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

likely to become targets of dispute initiation. As expected, larger traders are also more likely to become involved in a dispute. Because all three explanatory variables are in logarithms, the results imply that the effect of an increase in trade or economic size on dispute propensity is larger for low levels of trade and small economies than for large traders and large economies.

To further assess the gravity mechanism, we examine whether economic size affects dispute propensity through greater export diversity and re-estimate the gravity model (and the full models, as shown in Table 3) using the complainant's export diversity instead of its economic size. The results (available upon request from the authors) confirm that export diversity has a strong and statistically significant effect on trade dispute propensity. The results for the other variables remain essentially the same. Additional analyses confirmed that country size and export diversity are highly correlated, thus completing the causal chain from complainant's economic size to export diversity to trade conflicts. Hence we conclude that gravity effects are empirically well represented by including economic size and trade volume indicators in the models, without having to account explicitly also for trade diversity per se.

The results for the standard power and the preponderance models in columns 2 and 3 differ considerably. The model in the second column includes the relative economic size and relative income variables. This specification reflects the idea that it matters whether the complainant or the defendant is the more powerful country, and it corresponds to previously analysed power arguments in the WTO disputes literature (e.g., Guzman & Simmons 2005). The results clearly reject the hypothesis that this form of power matters. Both the relative economic size and relative income variables are statistically insignificant and do not increase the explanatory power of the model. Compared to the gravity model, the Bayesian Information Criterion (BIC) for the standard power model increases by more than 19 points, indicating that the former fits the data better than the latter. Similarly, the dispersion parameter, $\hat{\alpha}$, increases from 1.65 for the gravity model to 2.58 for the standard power model. The gravity model thus explains much more of the unexplained heterogeneity that generates variation in expected disputes across WTO dyads.

The power preponderance model in the third column produces different results. While income asymmetry does not have a statistically significant effect, economic size asymmetry is an important determinant of dispute propensity. The estimated coefficient indicates that, as expected, greater power asymmetry decreases the probability that a dyad experiences a WTO dispute. This result suggests that larger economies, both as complainants and defendants, may impose their will on smaller countries outside the WTO, which reduces dispute propensity as power asymmetry grows. The BIC indicates that the model fit for the preponderance model is almost as good as for the gravity model. The

statistically insignificant effect of income asymmetry suggests that specifying power asymmetry in terms of economic size differences is more appropriate. We conclude from these first tests of the two models that power asymmetry, but not relative power, plays a significant role in WTO trade dispute initiation.

The results for the legal capacity models, shown in the fourth and fifth columns, do not support the legal capacity hypothesis. The estimates imply that richer countries, as measured by GDP per capita, are less likely to initiate trade disputes and less likely to become targets of dispute initiation, after controlling for trade volume, economic size and other factors. We carried out several additional tests to assess the robustness of this finding. The latter remains the same for different specifications as long as we control for gravity effects – that is, whenever the model includes the log values of bilateral trade and/or export dependence and economic size. Only when we do not account for gravity do the results change and indicate that poorer countries become involved in trade conflicts less often.[13] Given the importance of gravity effects, we believe that the specifications, which include some or all gravity variables, are the most appropriate ones.

The capacity model in the fourth column uses a more specific measure of legal capacity: WTO delegation sizes of complainant and defendant countries. The effects of these two variables are not statistically significant.[14] A potential problem with the delegation size variable is that it may also capture gravity effects. Larger economies have a larger delegation because they are involved in more disputes. A country's delegation size, then, to a considerable extent, reflects its dispute propensity. Conversely, small economies have small or no delegations because they are rarely or never involved in trade conflicts.[15] This may explain why the impact of delegation size disappears when we include the proper gravitation variables in the model.

The results reported so far remain the same when we include all key variables in one model. The results for these full models are shown in columns 6 and 7 of Table 3. The results are also robust to estimating the models without dyads that include both the EU and the United States. Moreover, the results remain the same when we include the complainant's export diversity instead of its economic size.

To what extent are the results reported so far substantively relevant? Table 4 shows the predicted probabilities that *no* trade conflict occurs, using the gravity model and the best-fitting power preponderance and capacity models (columns 3 and 4 of Table 3) for different levels of trade. The first row for each model in Table 4 shows the predicted probabilities from these models for a reference dyad, which is a dyad with all explanatory variables kept at their means of the subset of dyads with at least one conflict.[16] The second and third rows for the gravity (capacity) model show the predicted probabilities

*Table 4.*  Predicted probabilities of no conflict for standardised dyads at different trade levels

| Trade | Mean | High | Maximum |
|---|---|---|---|
| *Gravity model* | | | |
| Reference dyad | 0.96 (0.95; 0.97) | 0.9 (0.85; 0.94) | 0.76 (0.61; 0.91) |
| Large size countries | 0.88 (0.81; 0.96) | 0.74 (0.66; 0.82) | 0.53 (0.46; 0.61) |
| Maximum size countries | 0.83 (0.68; 0.98) | 0.66 (0.51; 0.81) | 0.44 (0.34; 0.53) |
| *Power preponderance model* | | | |
| Reference dyad | 0.94 (0.92; 0.96) | 0.77 (0.70; 0.83) | 0.45 (0.37; 0.54) |
| Large asymmetry | 0.97 (0.95; 0.98) | 0.85 (0.80; 0.89) | 0.56 (0.47; 0.65) |
| Maximum asymmetry | 0.98 (0.97; 0.99) | 0.91 (0.85; 0.97) | 0.67 (0.53; 0.82) |
| *Capacity model* | | | |
| Reference dyad | 0.96 (0.94; 0.97) | 0.79 (0.74; 0.83) | 0.46 (0.38; 0.54) |
| Large income countries | 0.97 (0.97; 0.98) | 0.86 (0.83; 0.89) | 0.56 (0.49; 0.63) |
| Maximum income countries | 0.98 (0.97; 0.99) | 0.88 (0.84; 0.91) | 0.59 (0.51; 0.66) |

Notes:  Predicted probabilities are shown for different trade levels because marginal effects differ depending on trade. Mean/High/Maximum trade suggests that Log(Trade) is kept at the mean of the subset of dyads with at least one conflict/one standard deviation above the mean/at the maximum. Reference dyad means that the independent variables (country sizes, asymmetries and income levels) are at the mean of the subset of dyads with at least one conflict. For large size/asymmetry/income dyads, the respective explanatory variables are set one standard deviation above their means. For maximum size/asymmetry/income dyads, the respective variables are set at the maximum. 95 per cent confidence intervals are computed using the delta method and are listed in parentheses.

when the complainant's and the defendant's economic size (income) increase by one standard deviation or to the maximum. The same rows for the power preponderance model show the probabilities when the complainant's economic size increases and the defendant's size decreases by one standard deviation or to the maximum/minimum.

The effects are largest for the gravity model, and greater for the power preponderance model than for the capacity model. For the gravity model, the predicted probability of no conflict decreases by 8 (13) percentage points when we move from the reference dyad to the large (maximum) size dyad for a mean level of trade. The magnitude of the marginal effect differs depending on the volume of bilateral trade. The predicted probability of no conflict decreases by 16 and 24 percentage points, respectively, for a high level of trade and by 23 and 32 percentage points, respectively, at the maximum level of trade. For ease of interpretation, the mean trade level corresponds approximately to the trade level of dyads like the Czech Republic and Poland or Australia and Thailand. Trade between Argentina and Brazil is slightly

more than half way between mean and high trade, while, for example, Malaysia and Singapore or South Korea and Japan fall approximately into the high-trade category.

When power asymmetry increases by the same standardised amounts, the effect of moving to the maximum asymmetry dyad is only four percentage points for the mean value of trade, but this effect increases to 22 percentage points for maximum trade. It is important to note that the dispute-reducing effect of asymmetry is only visible at larger levels of trade because at low levels the probability of a conflict is close to zero and therefore cannot decrease further. In comparison, changes in income have a smaller substantive impact with marginal effects that vary between two percentage points for the mean trade dyad up to 13 percentage points for a dyad with maximum trade.

Do these results remain the same when we pay greater attention to dynamics and use a more sophisticated statistical method to deal with the rare events problem? To that end, we re-examine the best-fitting model in Table 3 using time-series cross-section data and zero-inflated count models. This approach distinguishes between dispute-enabling and dispute-promoting factors. The zero-inflated models are very sensitive to improper specifications. To keep the model simple and tractable, we only include variables in the equation where we expect their primary effect. We mainly use theoretical considerations to identify an appropriate empirical specification for the zero-inflated model, but also rely on empirical tests and results.[17]

In principle, legal capacity primarily should be dispute-enabling, implying that the capacity indicators should be included in the inflation equation, which estimates the probability that a conflict *never* occurs. Countries with no or low legal capacity simply are unable to participate in the WTO dispute settlement mechanism and thus should have a zero-probability of disputes. When the ability of countries to engage in WTO disputes increases because of greater legal capacity, the probability that a conflict never occurs (as estimated by the inflation equation) decreases. After a country reaches a certain level of legal capacity, the conflict propensity due to legal capacity should change only marginally. However, the model still distinguishes between a country with intermediate legal capacity that must carefully choose which cases it litigates and one with high legal capacity that can litigate as many cases as it wants.

We treat economic size and power asymmetry as primarily dispute-promoting. According to the logic of the gravity model, being a small country does not mean that a dispute will never occur, as estimated by the inflation equation. The same applies to the power preponderance argument, which does not say that small countries never litigate against a large country. As we can see in the data, small countries in fact have participated in the dispute settlement mechanism and also have initiated disputes against large countries. The

*Table 5.* Panel (zero-inflated negative binomial) model for full specification

|  | Full model (zero-inflated) | |
|  | Inflation equation | Conflict equation |
| --- | --- | --- |
| Log(Trade) | –0.487*** (–0.093) | 0.096 (–0.133) |
| Log(GDP A) |  | 0.288*** (–0.078) |
| Log(GDP B) |  | 0.364*** (–0.074) |
| Power Asym. |  | –0.231*** (–0.06) |
| Retaliation | –0.215 (–0.417) | 0.690*** (–0.142) |
| Log(Income A) | 0.198 (–0.158) |  |
| Log(Income B) | 0.18 (–0.133) |  |
| Democracy A | –0.182*** (–0.045) |  |
| Democracy B | –0.148*** (–0.042) |  |
| Constant | 3.452*** (–1.071) | –1.432 (–0.914) |
| $\hat{\alpha}$ | 0.25 | |
| N (all/non-zero) | 63,833/230 | |
| $\chi^2$ | 120.65 | |
| Prob > $\chi^2$ | 0.000 | |
| BIC | 2266.48 | |
| Log likelihood | –1055.77 | |

Notes: The inflation equation estimates the probability that a dyad is never involved in a trade dispute (belongs to the Always 0 group). The count equation estimates the expected number of trade disputes, weighted by the probability that a dyad belongs to the Not Always 0 group. Robust standard errors cluster on dyads and are listed in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

arguments imply, however, that we should observe a lower number of disputes in a dyad if a country is small or if the power asymmetry is large. This effect is captured by the count equation.

Table 5 shows the results for the zero-inflated model. It includes the same variables as the model in the last column of Table 3, plus the retaliation variable. Overall, the results shown in Table 5 confirm the earlier findings.[18] A notable result, which increases the confidence in our findings overall, is the large decrease of the dispersion parameter $\hat{\alpha}$. This parameter drops very close to zero, which means that a standard count (Poisson) model may also be used.[19] Re-estimating the model with a zero-inflated Poisson model yields almost identical results. We conclude from these results that much of the over-dispersion in the models in Table 3 is due to the large number of zeros (no dispute initiations) in the dataset.

As expected, economic size of complainants and defendants, power asymmetry and retaliation are important dispute-promoting factors. The estimated

coefficients of the size and power asymmetry variables in the conflict equation, the equation that estimates the expected number of disputes, are very similar to the ones in Table 3. The results are different for the capacity variables. We include these variables in the inflation equation. This choice is motivated by the idea that, if legal capacity has any impact, it should be a dispute-enabling factor. Unlike in the cross-sectional analyses, the effect of a country's income is not statistically significant. The results are the same for a specification including the number of delegates as a proxy for legal capacity. The zero-inflated model thus offers no support for the capacity hypothesis.

The results also suggest that trade is only dispute enabling; more trade does not have a dispute promoting effect – the coefficient on the log of trade is only statistically significant in the inflation equation, but not in the conflict equation. This finding supports our conceptual distinction between trade as the equivalent of distance in the classic gravity model and economic size as representing economic diversity and market size. If trade mainly represented diversity, we should observe an important dispute promoting effect – that is, the coefficient on trade volume in the conflict equation should be statistically significant and positive, which is not the case. Overall, we conclude that the dispute-promoting effect of trade is considerably smaller than suggested by the models in Table 4.

Table 6 illustrates the substantive gravitation and power effects predicted by the zero-inflated model.[20] The explanatory variables for the reference dyad are set to the same values as in Table 4. The large economic size (large asymmetry) dyad shows the predicted probabilities when complainant and defendant size (power asymmetry) increase by one standard deviation. The first row for each dyad [i.e., Pr(Always 0)] indicates the estimated probability from the inflation equation that a dispute will never occur in this dyad. The second row [i.e., Pr(0)] shows the overall probability from both equations that no conflict will occur. The predicted probabilities are computed for different levels of trade. Mean, high and maximal trade are as defined in Table 4.

The predicted probabilities confirm the results from the standard count model, but the magnitudes change and are more realistic. The probability of being in the Always 0 group is 39 per cent for a mean trade level (first column), which is rather low because the democracy and the trade levels of countries with at least one dispute are already fairly high.[21] When trade increases, the probability of being in the Always 0 group converges towards zero because of the dispute-enabling effect of trade (third column). The expected actual number of conflicts changes primarily when conflict-promoting factors like country size and power asymmetry vary. The probability of experiencing no conflict decreases by up to 30 per cent when country size increases by one standard deviation.[22] Similarly, the same probability increases by up to 23 per cent when power asymmetry increases by the same amount.

*Table 6.* Predicted probabilities of *no* conflict for zero-inflated model

| Trade | | Mean | High | Maximum |
|---|---|---|---|---|
| Reference dyad | Pr(Always 0) | 0.39 (0.04; 0.63) | 0.15 (0.01; 0.32) | 0.08 (0.01; 0.19) |
| | Pr(0) | 0.73 (0.63; 0.84) | 0.56 (0.42; 0.72) | 0.48 (0.30; 0.68) |
| Large size countries (Gravity effect) | Pr(Always 0) | 0.39 (0.04; 0.63) | 0.15 (0.01; 0.32) | 0.08 (0.01; 0.19) |
| | Pr(0) | 0.5 (0.35; 0.67) | 0.26 (0.17; 0.39) | 0.18 (0.11; 0.29) |
| Large asymmetry (Power effect) | Pr(Always 0) | 0.39 (0.04; 0.63) | 0.15 (0.01; 0.32) | 0.08 (0.01; 0.19) |
| | Pr(0) | 0.86 (0.79; 0.91) | 0.76 (0.65; 0.84) | 0.71 (0.57; 0.83) |

Notes: Pr(Always 0) is the probability that the dyad does not have any chance to be involved in a trade dispute (i.e., that the dyad is in the Always 0 group). This probability is estimated using the inflation equation of the model in Table 5. Pr(0) is the overall probability that no WTO trade conflict occurs. This probability is estimated using both the inflation and count equations of the model in Table 5. Reference dyads and large size/asymmetry are defined as in Table 3. The numbers in parentheses are 95 per cent confidence intervals that were computed using bootstrapping. The bootstrapping results are based on 1,000 replications.

## Conclusion

This article contributes to clarifying whether variation in participation of countries in WTO dispute settlement is driven mainly by gravitational forces emanating from economic size and trade volume, or whether discriminatory power and capacity biases are also at work. We have placed this analysis in the larger context of international relations research on legalisation beyond the nation-state. Unlike previous research, we examine different possible mechanisms of power biases and also distinguish between dispute-enabling and dispute-promoting effects of gravity, power and capacity variables.

The analysis shows that the strong presence of large industrialised countries and the absence of least developed countries in WTO dispute settlement is caused primarily by gravitation, but to some extent also by power preponderance. Larger economies, because of their greater economic diversity and market size, gravitate towards trade disputes to a greater extent than their larger trade volume alone predicts. In addition, country pairs characterised by a large power asymmetry are less likely to experience a WTO trade dispute. The latter finding suggests that trade disputes in such dyads may be dealt with outside the WTO.

By and large, our findings can be viewed as moderately good news for proponents of legalisation in the international system. While it is obvious that legalisation has advanced at varying speeds in different areas of international relations, the evidence for the WTO system shows that considerable progress has been made in at least one important realm. The finding that, controlling for other factors, larger economies and bigger traders become involved in more trade disputes (the gravitational effect) hardly can be interpreted, from our viewpoint, as discrimination. The fact that we were not able to identify discriminatory effects pertaining to legal capacity is good news as well, even though further research based on improved indicators of legal capacity and a better handling of endogeneity problems is needed.

Our results concerning power asymmetry (preponderance effect) are potentially worrying. It is, arguably, much easier to reduce legal capacity differences than to reduce power differences. However, additional research (e.g., case studies on a larger scale) is required to establish the conditions under which smaller and/or poorer countries experiencing costly trade restrictions are, due to power asymmetry, driven into concessions outside the WTO system.

## Acknowledgements

Political Economy of International Organisations (PEIO) conference in Geneva. We thank Todd Allee, Xun Cao, Christina Davis, Alexander Keck, Stephanie Rickard, Johannes Urpelainen and the other participants at these conferences for their valuable comments, as well as Todd Allee, Manfred Elsig and Henrik Horn for providing data.

## Notes

1. Theoretically, it is possible for a country to have large GDP, but a low degree of economic diversity. The gravity argument would not apply to such a country. However, countries of this type are very rare. Exceptions are those with large natural resource endowments (e.g., oil-exporting countries). We conducted robustness checks, excluding these countries. The results are robust to including or excluding such countries. We also examined whether the assumption that larger economies are economically more diverse is empirically plausible (see results section).

2. Variation in trade volume may also indicate differences in trade diversity that emanate from the economic size of countries, but only to a limited degree. E.g., small economies that are close to each other are likely to trade more than distant countries, but trade between these two small countries is not necessarily very diverse. For now, we keep the conceptual distinction between trade reflecting distance and economic size representing diversity, but we will return to this issue when we examine the distinction between conflict-enabling and conflict-promoting factors.

3. On irrelevant dyads in interstate war research, see Russett and Oneal (1997), and Lemke and Reed (2001). Since trade conflicts are next to impossible if countries do not trade, non-trading dyads are considered irrelevant. We keep all dyads with a positive bilateral trade volume.

4. Alternatively, one could choose a design that first assesses the conflict potential of a dyad. The advantage of our design is that we are able to analyse *all* potential conflict dyads. We believe that this approach is particularly useful for an analysis of power and capacity arguments because we can cover the complete range of the power distribution within the WTO. Research choosing the alternative approach (e.g., Allee 2004; Bown 2005a; Davis & Shirato 2007; Kim 2008) is less encompassing than ours because it is usually restricted in the scope of sectors, countries or types of trade barriers covered in the analysis.

5. We use data by Kim (2008) on antidumping duties or counterveiling duties (AD/CVD) that one country imposes on its trading partners. The AD/CVD indicator measures the potential of a trade conflict because it reflects whether a potentially WTO-inconsistent trade policy that could be challenged by a trading partner exists.

6. The choice of statistical methods is discussed in greater detail below.

7. See www.wto.org/english/tratop_e/dispu_e/dispu_e.htm

8. These data are based on the WTO telephone directory in October 2002 (see also http://econ.worldbank.org/).

9. Whether a dispute is initiated may also depend on the characteristics of the economic sector concerned. The dyadic approach followed here, which is particularly useful for the analysis of gravity, power and capacity effects, does not allow for a sectoral explanation of disputes because the dependent variable cannot distinguish between sectors. We believe that the omission of sectoral characteristics does not bias our estimates because there is no

*a priori* reason why these should be correlated with our indicators for gravity, power and capacity. For studies on sectoral effects, see Allee (2003, 2004); Davis & Shirato (2007).

10. In the cross-sectional dataset, the maximum number of disputes in a dyad is 30 (USA–EU).

11. Using the average number of disputes per year would be equivalent to using the total number of conflicts as long as the exposure time (i.e., the start date of WTO membership) was the same across countries. Apart from the variable scale, the difference between the total number and the average number of conflicts is that the latter weights the former by the number of years of WTO membership, which is the same for the vast majority of observations. As discussed in the main text, the cross-section analysis explicitly accounts for differences in exposure time as measured in terms of number of years when a WTO dispute can happen.

12. Relevant means that a dyad has a positive probability of a trade conflict. For theoretical reasons, we opt for the zero-inflated model instead of the related hurdle model (King 1989). The latter assumes that an event always occurs when a specific threshold is crossed. This condition is not met in our empirical area. In the zero-inflated model, some dyads have a zero probability of becoming involved in a trade dispute – e.g., because they trade only very little (the Always 0 group). Dyads that trade more may be subject to a non-zero probability of a dispute, but this does not mean that a trade dispute will necessarily occur (the Not Always 0 group). In practice, results from the two models are often very similar (Zorn 1998).

13. In an alternative specification, which includes the gravity variables without the logarithmic transformation, GDP per capita has the expected positive effect on dispute initiation. All previously discussed results concerning gravitation, relative power and power asymmetry hold for specifications with non-log variables. We opt for the logarithmic transformation, for two reasons. First, the idea that marginal effects at low levels of trade, economic size and income are larger than at high levels of these variables is theoretically plausible. Second, empirical tests clearly show that the specifications with the log-transformation are more appropriate.

14. As for GDP per capita, the effect of delegation size in specifications with non-log gravity variables, or one that does not capture gravitational effects at all, changes. The results then imply that greater legal capacity increases dispute propensity. We opt for the specification with the log-transformation for reasons discussed in the previous note.

15. Accordingly, the log of GDP is highly correlated with the delegation size of a country.

16. This choice is motivated by the fact that trade disputes are very rare events and probabilities at the mean levels of the full sample are always very close to zero – e.g., because most country dyads in our dataset trade very little. We would not be able to judge the dispute-reducing effect of power asymmetry adequately for dyads with a zero probability of a dispute.

17. The MLE estimator of zero-inflated negative binomial models often fails to converge if the models are not properly specified – i.e., if they include many variables that are not relevant for an equation. We therefore impose a number of assumptions about the data-generating process that we develop based on theoretical considerations (dispute-enabling versus dispute-promoting factors) and empirical tests. While these assumptions can be challenged and improved through further research, the good performance and plausible results of our model suggest that these assumptions are reasonable.

18. Note that the signs of the coefficients for the variables that are included in both equations should point in opposite directions across equations.

19. A likelihood ratio test suggests that the null hypothesis that α equals zero cannot be rejected at the 1 per cent significance level.
20. We do not report capacity effects because the models show that there are no significant effects of this kind.
21. Setting trade to a low level and taking into account that many countries are not democratic, the same probability increases to above 90 per cent, which is plausible for most dyads.
22. Gravity effects are computed by subtracting Pr(0) for large country size (second row) from Pr(0) for the reference dyad (first row) for the different trade levels (columns 1–3). Power effects are computed accordingly.

## References

Abbott, K.W. et al. (2000). The concept of legalization. *International Organization* 54(3): 401–419.

Allee, T. (2003). Going to Geneva? Trade Protection and Dispute Resolution under the GATT and WTO. Manuscript. University of Illinois.

Allee, T. (2004). Legal Incentives and Domestic Rewards: The Selection of Trade Disputes for GATT/WTO Dispute Resolution. Manuscript. University of Illinois.

Allee, T. (2005). Developing Countries and the Initiation of GATT/WTO Disputes. Paper presented at the annual meeting of the International Studies Association. Honolulu, HI.

Bergstrand, J.H. (1985). The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *Review of Economics and Statistics* 67(3): 474–481.

Bernauer, T. & Sattler, T. (2006). Sind WTO-Konflikte im Umwelt- und Verbraucherschutz eskalationsträchtiger als andere WTO Konflikte? *Zeitschrift für Internationale Beziehungen* 13(1): 1–25.

Bown, C.P. (2005a). Participation in WTO dispute settlement: Complainants, interested parties and free riders. *World Bank Economic Review* 19(2): 287–310.

Bown, C.P. (2005b). Trade remedies and World Trade Organization dispute settlement: Why are so few challenged? *Journal of Legal Studies* 34(2): 515–555.

Busch, M.L. (2000). Democracy, consultation and the paneling of disputes under GATT. *Journal of Conflict Resolution* 44(4): 425–446.

Busch, M.L., Reinhardt, E. & Shaffer, G. (2008). *Does legal capacity matter? Explaining dispute initiation and antidumping actions in the WTO*. Issue paper 4. Geneva: International Centre for Trade and Sustainable Development.

Conybeare, J. (1985). Trade wars: A comparative study of Anglo-Hanse, Franco-Italian and Hawley-Smoot conflicts. *World Politics* 38(1): 147–172.

Davis, C.L. & Blodgett Bermeo, S. (2009). Who files? Developing country participation in WTO adjudication. *Journal of Politics* 71(3): 1033–1049.

Davis, C.L. & Shirato, Y. (2007). Firms, governments and WTO adjudication: Japan's selection of WTO disputes. *World Politics* 59(2): 274–313.

Feenstra, R.C. et al. (2005). *World trade flows, 1961–2000*. Working paper 11040. Cambridge, MA: National Bureau of Economic Research.

Gleditsch, K.S. (2002). Expanded trade and GDP data. *Journal of Conflict Resolution* 46(5): 712–724.

Goldstein, J. et al. (2000). Introduction: Legalization and world politics. *International Organization* 54(3): 385–399.

Greene, W.H. (1994). *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models*. Working paper 94–10. New York: Stern School of Business.

Guzman, A. & Simmons, B.A. (2005). Power plays and capacity constraints: The selection of defendants in WTO disputes. *Journal of Legal Studies* 34(2): 557–598.

Hegre, H. (2008). Gravitating toward war: Preponderance may pacify, but power kills. *Journal of Conflict Resolution* 52(4): 566–589.

Horn, H. & Mavroidis, P.C. (2008). *The WTO dispute settlement system, 2005–2006: Some descriptive statistics*. Working paper 740. Stockholm: Research Institute of Industrial Economics (IFN).

Horn, H., Mavroidis, P.C. & Nordström, H. (1999). *Is the use of the WTO dispute settlement system biased?* Discussion paper 2340. London: Centre for Economic Policy Research.

Kim, M. (2008). Costly procedures: Divergent effects of legalization in the GATT/WTO dispute settlement procedures. *International Studies Quarterly* 52(3): 657–686.

King, G. (1988). Statistical models for political science event counts: Bias in conventional procedures and evidence for the exponential Poisson regression model. *American Journal of Political Science* 32(3): 838–863.

King, G. (1989). Event count models for international relations: Generalizations and applications. *International Studies Quarterly* 33(2): 123–147.

Lemke, D. & Reed, W. (2001). The relevance of politically relevant dyads. *Journal of Conflict Resolution* 45(1): 126–144.

Marshall, M.G., Jaggers, K. & Gurr, T. (2002). *Polity IV project: Political regime characteristics and transitions, 1800–2003*. Available online at: www.cidcm.umd.edu/inscr/polity/.

Reinhardt, E. (1999). Aggressive Multilateralism: The Determinants of GATT/WTO Dispute Initiation, 1948–1998. Manuscript. Emory University.

Rosendorff, B.P. (2005). Stability and rigidity: The dispute resolution mechanism at the WTO. *American Political Science Review* 99(3): 389–400.

Russett, B.M. & Oneal, J.R. (1997). The classical liberals were tight: Democracy, interdependence and conflict, 1950–85. *International Studies Quarterly* 41(2): 267–294.

Smith, J. (2004). Inequality in international trade? Developing countries and institutional change in WTO dispute settlement. *Review of International Political Economy* 11(3): 542–573.

World Trade Organisation (WTO). (2006). *World trade report*. Available online at: www.wto.org/english/res_e/reser_e/wtr_e.htm

Zangl, B. (2008). Judicialization matters! A comparison of dispute settlement under GATT and the WTO. *International Studies Quarterly* 52(4): 825–854.

Zorn, C.J.W. (1998). An analytic and empirical examination of zero-inflated and hurdle Poisson specifications. *Sociological Methods and Research* 26(3): 368–400.

*Address for correspondence*: Thomas Sattler, School of Politics and International Relations, University College Dublin, Belfield, Dublin 4, Ireland. E-mail: thomas.sattler@ucd.ie