IFN Working Paper No. 1219, 2018

Efficiency Gains and Time-savings of Permanent Panels in the WTO Dispute Settlement

Louise Johannesson

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Efficiency gains and time-savings of permanent panels in the WTO dispute settlement

Louise Johannesson*

1 Introduction

Although the success of the dispute settlement mechanism (DSM) of the World Trade Organization (WTO) has been received enthusiastically, it has also brought on new challenges that have proved difficult to resolve. Issues such as increasingly time-consuming and complex disputes have strained the capacity of the WTO and led to delays. The WTO has long prided itself on prompt dispute resolutions, recognizing that excessive delays may force countries to find alternative solutions, which would, in the long-run, undermine the legitimacy of the DSM. Recent increases in the number of appeals (88% of all panel reports were appealed in 2016) prompted the current chairman of the Appellate Body (AB) Ujal Singh Bhatia, to explicitly address the problem of delays in the DSM.[1] He explained the situation and the consequences if the problem remains unresolved as: "When delays in WTO dispute resolution become the norm, they cast doubt on the value of the WTO's rules-oriented system itself. An erosion of trust in this system can lead to the re-emergence of power orientation in international trade policy". Apart from legitimacy concerns, timeliness is also a serious economic issue since delays in rulings also delay the removal of the

*Research Institute of Industrial Economics, Box 55665, SE-102 15 Stockholm and Örebro University, Sweden. E-mail: louise.johannesson@ifn.se. For valuable comments, I would like to express my gratitude to Henrik Horn at IFN, Magnus Lodefalk and Dan Johansson at Örebro University, Hildegunn Kyvik Nordås at the OECD, and Jonas Björnerstedt at Södertörn University. I gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation and the Marianne and Marcus Wallenberg Foundation.

[1]The event in question was a presentation of the Appellate Body's 2016 Annual Report. https://www.wto.org/english/news_e/news17_e/ab_08jun17_e.htm

illegal trade barriers. And the DSM does not award retro-active compensation in case of a violation.

Canada voiced concern over delays in a Dispute Settlement Body (DSB) meeting[2] in relation to dispute DS 384, stating that they had faced numerous delays that had unduly prolonged the economic burden on the Canadian economy. Korea expressed a similar sentiment in 2015[3] when they acted as a complainant in a dispute concerning anti-dumping measures by the US against certain Korean steel products (DS 488). They argued that excessive delays resulted in losses of ten million US dollars a month for Korean firms, which would cause a substantial number of job losses and failed businesses. The statutory deadline of panel proceedings is six months, but the average length of a panel process is around 15 months. This implies that an additional nine months are added to the proceedings, and in the case of the Korean anti-dumping dispute, an additional loss of 90 million US dollars. They further expounded on the detrimental consequences of protracted adjudications: "Long delays created perverse incentives by lowering the cost of adopting and maintaining WTO-inconsistent measures. Interest groups seeking protection would pressure Members to adopt these measures, insisting rightly, that they would not be subject to review by the WTO for years. Members could therefore expect more protectionist measures, and more, not less, disputes being brought to the WTO". In summary, it is evident that undue delays cause economic harm and that the DSM will be weakened if the problem is allowed to persist, forcing countries to seek alternative solutions for their trade skirmishes, alternatives that may be less transparent and based more on negotiating power.

The capacity constraint of the DSM clearly requires a comprehensive approach, where all the different stages are investigated to improve overall timeliness. One reform in that direction may be to install permanent panels. The idea of permanent panels, instead of the current regime of *ad hoc* panels, has been a recurring subject since panels were introduced into the GATT (Jackson, 1979; Tuthill et al., 1985). Various WTO members, most notably the EU (2002)[4] and academics, such as Hudec (1999), Davey (2003), Bourgeois (2003), Cottier (2003) and Busch and Pelc (2009), have expressed support for some variant of a permanent panel, as it will automatically save time by eliminating the, at times protrac-

[2] WT/DSB/M/362
[3] WT/DSB/M/367
[4] TN/DS/W/1

ted selection process. Furthermore, it has also been argued that a permanent panel could increase efficiency and quality in rulings as panelists will be more experienced than *ad hoc* panelists (for example Busch and Pelc, 2009). However, political apprehension, especially from the US (Cottier, 2003) has effectively made such proposals moot. As Bourgeois (2003) notes, "the underlying tension lies between legitimacy and efficiency", where scholars and trade experts give priority to the latter, while the politicians stress the former, and for the WTO, an inter-governmental body, the weight of legitimacy lies heavy in the debate. However, if the problem of delays becomes a bigger and bigger threat to the legitimacy of the DSM, there may be cause to re-consider the importance of efficiency vis-à-vis legitimacy. This paper contributes to this discussion by analyzing certain panel features that characterize a permanent panel, namely, the panelists' experience and whether a prior work relationship between the panelists facilitates their deliberations. We examine how these factors contribute to the time it takes to issue a panel report, as used in previous studies, but also a panel's efficiency in examining claims.

The results show that the combined experience of the panel does not reduce the time it takes to issue a report, as found in Horn (2013), but it slightly decreases efficiency. One reason why more experience leads to longer panel deliberations could be selection effects. That is, more experienced panelists are appointed in more complicated cases, which in turn takes longer to complete. We find, however, that panelists that work with previous co-panelists can reduce the time by around 30 days, a result that weighs in favour of permanent panels.

The rest of the paper is organized as follows. Section 2 discusses the arguments for and against permanent panels and the expected consequences. Section 3 describes the data, variables and method, concluding with descriptive statistics. Section 4 characterizes the institutional setting in which the disputes are examined. Section 5 presents the results of the analysis, with a short comment on robustness. Section 6 concludes with a discussion.

2 Permanent or *ad hoc* panels

A constitutive argument against permanent panels has been that an *ad hoc* panel embodies the spirit of the WTO as a member-driven organisation. Nonetheless, it has not gone unnoticed that the majority, around 66% of all panels between 1995–2014, have actually

been composed by the Director-General (DG) together with the Secretariat, and not by the parties to the dispute; a trend that seem to have grown over time. One reason that panel composition has been increasingly referred to the DG could be related to increased complexity. Horn (2013) examines the determinants of time required to issue panel reports, and he shows that panels that were appointed by the DG take longer to complete a panel report compared to panels that were composed by the parties. This may indicate that the DG is more often consulted in politically sensitive disputes.

One of the strongest arguments in favour of a permanent panel is that more experienced panelists will increase the quality of the panel report, which will reduce the number of reversals by the AB. Busch and Pelc (2009) examined this claim and found that, only the experience of the chair significantly reduced the probability of the AB overturning a ruling. This supports a type of mixed panel suggested by Cottier (2003), with permanent chairs and *ad hoc* panelists. Horn (2013) also finds that the experience of panelists does not facilitate the timeliness of a panel report, a result that further weakens the argument for permanent panels. Additionally, he finds that a larger body of case law, which should simplify the work, had no impact on time.

Another argument against permanent panels has been the lost opportunity for socialization for *ad hoc* panelists (Davey, 2003). The introduction of the DSM to hundreds of such panelists may prove productive in the long-term for the reputation of the system. However, the flip-side of this argument, and a feature of permanent panels, is the fact that the same people will work together repeatedly and form closer work relationships compared to the panelists that are selected *ad hoc*. It is entirely possible that such an arrangement would improve work flow and save time.

3 Data and analysis

Data for this study is collected, mainly, from the publicly available "WTO dispute settlement database", originally compiled by Henrik Horn and Petros C. Mavroidis. It includes all initiated disputes from 1995 to 2015.[5] For descriptive statistics, we also use an institutional variable—annual number of regular staff at the WTO Secretariat—that is found in the

[5]The dataset can be freely downloaded here: http://globalgovernanceprogramme.eui.eu/wto-case-law-project

WTO Annual Reports (1998–2015), subsection "Secretariat and the Budget".

3.1 Variable description

The first outcome variable, time required to issue panel reports (hereafter "Duration") was calculated from the date of constitution of the panel to the date the panel circulated the report. This includes the interim period, where the panel circulates the panel report only to the affected parties. The statutory deadline is set at six months but Table 1 shows that the average duration is usually over a year. Although duration may seem trivially correlated with the number of claims, the relationship is not necessarily linear. It is reasonable to assume that there are economies of scale in the legal analysis; that is, examining two claims in a dispute does not take twice as long as examining a single claim. Panelists invest considerable amount of effort for the first few claims as they familiarize themselves with the facts and circumstances of the case; however, after this initial research, they are able to complete the legal analysis for the subsequent claims more expeditiously.

For that reason, we also construct a second outcome, defined as days per invoked claim, but each claim in a dispute is assigned different weights using the square root. This measure is referred to as a square root equivalence scale. Let u be days between the constitution of the panel and circulation of the panel report and c_d be all cited claims in dispute d. We compute equivalised duration per claim (henceforth Eq. days per claim), x_d as:

$$x_d = \frac{u_d}{\sqrt{c_d}}$$

We interpret x_d as follows: As the number of claims increases, the amount of time spent on each claim decreases non-linearly. For example, a panel that has spent 100 days on four claims will spend 50 Eq. days per claim ($100/\sqrt{4} = 50$). One way to interpret this variable is to imagine a panel that spends 50 days on two related claims concurrently, taking advantage of economies of scale, instead of assuming the panelists spend 25 days on each claim separately. An alternative definition to consider is to divide duration with unique claims rather than all cited claims to better capture the true time spent on legal analysis. However, this is exactly the type of situation that equivalisation should account for. (See Appendix E for estimation results using alternative outcomes.)

The explanatory variables that we are particularly interested in are "experience", "in-

crowd", and to some extent also "case law", since a rich case law may compensate low experience to some degree. Experience is simply computed as the sum of all previous panels that the current panelists have served on. We also include a squared-transformed version. Along the same line, the variable "in-crowd" is the sum of previous panels that two or more panelists in dispute d have served together on. Although being former co-panelists may make it easier to work together and thus increase efficiency, it could also be entirely possible that these previous interactions were unproductive and instead impede the process. But since we can only observe previous interactions, we try to better capture positive effects of established work relationships by including an interaction variable between the DG and in-crowd. We hypothesize that the DG and Secretariat are better at selecting a panel that work well together based on insider knowledge, or at least avoiding panel compositions that lead to excessive discord.

Apart from these variables of interest, we also include several other variables that should be associated with duration, as shown in Horn (2013). To capture case law, let c_d^j be the number of times a specific claim j is invoked in dispute d. We define W_d^j as the sum of all previous claims j in dispute d, so that $W_1^j = 0$.

$$W_d^j = \sum_h^{d-1} c_h^j,$$

where $d \leq 2$ and h is earlier disputes (Horn, 2013). The lack of substantial effects by richer case law found in Horn (2013) may be a result of the confounder "complexity". If the DSM is increasingly used as a way to "complete" an incomplete contract (i.e. the WTO, see Horn et al. (2010)), the current case law may, by definition, consist mainly of legally complicated disputes, so that a rich case law becomes an indication of legal complexity, which in turn implies little effect on timeliness. For this reason, we interact case law with disputes with domestic instruments claims.

As in Horn (2013), we consider three types of case complexity: political, subject and legal complexity. We capture political complexity by including an indicator of whether the panel was composed by the DG or by the parties. If the parties are unable to compose a panel through mutual agreement, it may signify a politically sensitive dispute, adding to the overall complexity. We account for subject complexity through the variable "Expert",

which indicates whether the panel consulted external experts. Lastly, legal complexity is captured by controlling for the share of all claims that concern domestic instruments. These cases are known to be contentious as it is highly difficult to accurately identify legitimate regulatory measures from disguised protectionism (Howse, 2012). The relevant provisions include the SPS Agreement (Agreement on the Application of Sanitary and Phytosanitary Measures); the TBT agreement (Agreement on Technical Barriers to Trade); article 2, 5, and 7 of the SCM agreement (Agreement on Subsidies and Countervailing Measures); article III of the GATT (National Treatment); and article XX(b) and XX(g) of the GATT (General Exceptions).

We also include an indicator for panels that had at least one developing country judge in a dispute between a developing and an industrialized country. This is motivated by the Special and Differential (SDT) provision Art. 8.10 of the DSU. It states: "When a dispute is between a developing country Member and a developed country Member the panel shall, if the developing country Member so requests, include at least one panelist from a developing country Member." Developing countries may be concerned that panel deliberations, consisting only of judges from developed countries, do not adequately consider development issues in the analysis. Art. 8.10 is a way to ensure that such issues are voiced. Thus, we include this indicator to potentially capture a higher degree of dissent among the panelists, which could increase the time it takes to issue the panel report. Note that we denote this variable as "Art. 8.10 composition" just as a shorthand for panels that include at least one developing-country judge in disputes that involve one developing and one industrialized country. It does not indicate whether the developing country in question actually invoked Art. 8.10 since such information is private. Lastly, we include the number of third parties to control for additional argumentation that the panel needs to consider, DG-appointed panels and multiple panels, which indicate panels where the panelists worked on several panels concurrently.

3.2 OLS and median regression

The nature of legal cases involves a great deal of heterogeneity across disputes, which produces a rather skewed distribution with a few extreme values.[6] These could be potential outliers, something that complicates our estimation since OLS is quite sensitive to extreme values. Often when using a sample, removing outliers can be justified by assuming homogeneous effects in OLS and arguing that the extreme value is non-representative. But in our case, the data we use are, in fact, population data as the whole universe of WTO disputes is encompassed. In theory, this does not necessarily alter the way we handle outliers, but in practice, it becomes an awkward exercise to remove observations from the actual population for the sake of an assumption. Instead, we approach the estimation as follows: In addition to ordinary least squares (OLS) regression, we also use least absolute deviation regression, i.e., median regression, which is robust against extreme values. In that way, we avoid over-reliance on the outliers in case they are, in fact, true outliers. If we assume that we have population data, OLS will be better at characterizing the actual distribution, while the median regression only characterize the 50th percentile.

The use of population data, however, requires a different perspective on how we interpret the estimated coefficients. One approach is to assume that the realized disputes, so far, constitute a finite population drawn from an infinite population of disputes, sometimes called a superpopulation. In such a framework, the effects are predictive since we use the realized population data to model the data-generating process of the superpopulation. In that case, the regression coefficients will be the usual estimates of a statistical relationships, for the superpopulation that is, and the standard errors will be a measure of statistical uncertainty. However, even though the concept of a superpopulation is reasonable, it is still a purely imagined population without real-world context, and for this reason, we also wish to describe the population itself as a historical account of past disputes. In this framework sampling variation is, naturally, zero, and any unaccounted variation in the statistics would stem from other sources, like measurement errors. The regression coefficients are viewed as average responses to marginal changes in the independent variable and the standard errors as approximate estimates of the normal standard deviation from the mean. (Since

[6] Most notably, the first and second Airbus-Boeing disputes (DS 316, 353) that took over four years to complete.

the sample standard error is adjusted only by $n-2$ rather than n to account for degrees of freedom in sample estimation.) We will, hence, estimate the following model:

$$y_d^k = \alpha + \mathrm{z}_d'\boldsymbol{\beta} + \gamma W_d^j + \varepsilon_d,$$

where y_{ds}^k is the dependent variable and $k = x_d, u_d$ indicates Eq. days per claim and duration, respectively. z_d is a vector of explanatory variables, and W_d^j is the available case law for dispute d.

3.3 Descriptive statistics

A first look at the data, presented in Table 1, shows that the average duration is little over 12 months (385 days) and the average Eq. days per claim is around 128. That is, for a dispute with 20 claims (the average number of claims), it will take approximately 1.6 years to issue a panel report if we use the average Eq. days per claim, or using the median, the expected duration is a little over 1 year (371 days), which is quite close to both the average mean and median duration (351 days). Compare this with using the unweighted days per claim of around two months, which would result in an estimated duration of around 3 years for 20 claims, or a little under two years (680 days) if we use the median. To put this in context, imagine that all disputes met the statutory deadline of 185 days; then, the panel would have to spend an average of 66 Eq. days per claim, or 33 unweighted days per claim. But, of course, whether 33 days per claim is better than 58 days per claim is another question.

The DG has appointed 70% of all panels, and according to Table 1, an average panel has a panel chair that has experience from 1–2 previous panel when they accept the position. (As of 2014, these panel chairs have overseen, on average, 5 panels.) The combined experience in a panel is around 5–6 previous panels when they are appointed and a panelist has, on average served on one previous panel with a current co-panelist. The panel can consult, on average, eleven previous judgments,[7] and around seven third parties join the proceedings, on average.

[7] The median is, however, much lower, at only four. The maximum value indicates a suspected outlier with a body of case law of 179. But removing this one observation only changes the average case law to 10.

Table 1: Descriptive statistics

Variables	Mean	SD	Median	Min	Max
Duration	385.3	190	351.0	126.0	1,717
Days per claim (equal weight)	58.8	64	34.5	2.4	316
Eq. days per claim	128.3	77	112.0	28.1	562
Interim days	87.2	45	77.0	28.0	299
No. of claims	19.9	24	11.0	1.0	135
No. of unique claims	8.8	7	6.0	1.0	37
In-crowd	1.2	2	0.0	0.0	8
Expert	0.1	0	0.0	0.0	1
Chair Experience	1.7	2	1.0	0.0	9
Total Panel Experience	5.8	5	5.0	0.0	25
Case law	11.4	19	4.0	0.0	179
Art. 8.10 composition	0.1	0	0.0	0.0	1
No. of Third Parties	6.9	5	6.0	0.0	24
Appointed by DG	0.7	0	1.0	0.0	1
Domestic instrument disputes	0.4	0	0.0	0.0	1

N = 210

Ten per cent of the panels were Art. 8.10 compositions. In Table 2, we see that both developing countries and BIC (Brazil, India and China) have mainly faced EU and the US in disputes. The size of the round shaded area represents the share of all disputes, while the number is the frequency. We see immediately that G2 is the most frequently targeted group, as well as the most active complainant. If we imagine this cross-tabulation as a coordinate plane, Quadrant I reveals that the majority of disputes are among the industrialized countries, while a very small share of disputes occur among developing countries (Quadrant III). The diagonal illustrates the disputes within the same development category. As expected, the most frequent interaction within the same group has been between the EU and the US.

Table 2: Complainant-respondent dyads, frequency and percentage

Complainant	BIC	DEV	G2	IND
IND	5	8	44	8
G2	23	12	26	26
DEV	1	12	26	2
BIC		1	26	3

Respondent

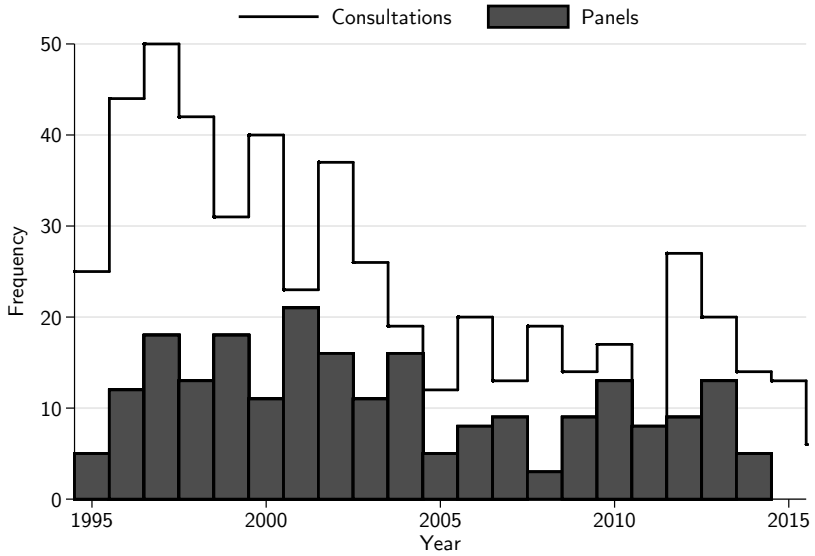
4 The institutional setting

As mentioned, this empirical study is based on a data set where the unit of observation is disputes. Since every dispute is unique, there is only variation between disputes and no variation within disputes over time. Since we are unable to estimate any institutional effects that influence duration and Eq. days per claim (unless we make bold assumptions.[8]), we offer a few institutional observations to consider alongside the dispute-level analysis.

Since the establishment of the WTO in 1995, over 500 disputes have been initiated and around 42% of these disputes have resulted in a panel report. Since all panels so far have

[8]Though there are methods, such as quasi-ML, that relaxes assumptions regarding normality, consistency would still be a concern due to our relatively small data set.

Figure 1: Frequency of consultations and panels



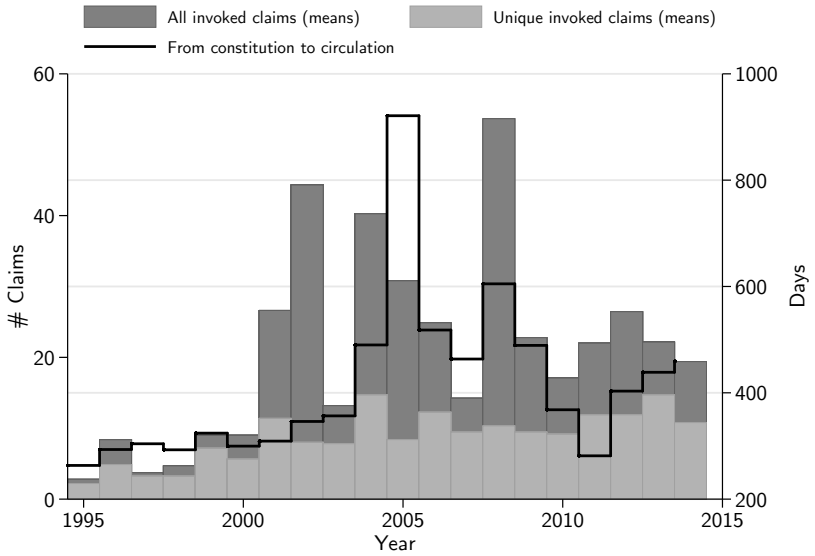
involved three panelists, it tallies to 669 appointed panelists over 223 panels.[9]

Figure 1 shows that the number of constituted panels has fallen slightly over the years but not as markedly as the decrease in the number of initiated disputes. This can, more or less, be explained the falling interest of the US and the EU (Appendix B). One reason for this trend may be that there were a great deal of low-hanging fruit in the beginning and as there are fewer and fewer "easy" disputes, there is a narrower selection of increasingly challenging disputes, which results in longer and longer panel proceedings.We try to account for this underlying complexity using various complexity measures as described above. A notable exception in this trend is 2012, when there was a sudden spike in the number of complaints. This coincided with a decrease in the number of senior Secretariat lawyers due to policy changes by the DG at the time, Pascal Lamy,[10] resulting in a sharp increase in the total time spent at the panel stage, as seen in Figure 2. In 2012, it jumped by over

[9] A total of 257 panels have officially been established, but not all of them result in a panel report.

[10] White & Case LLP, JDSupra.com (2014). Source:http://www.jdsupra.com/legalnews/wto-dispute-settlementlong-delays-hit-t-39550/

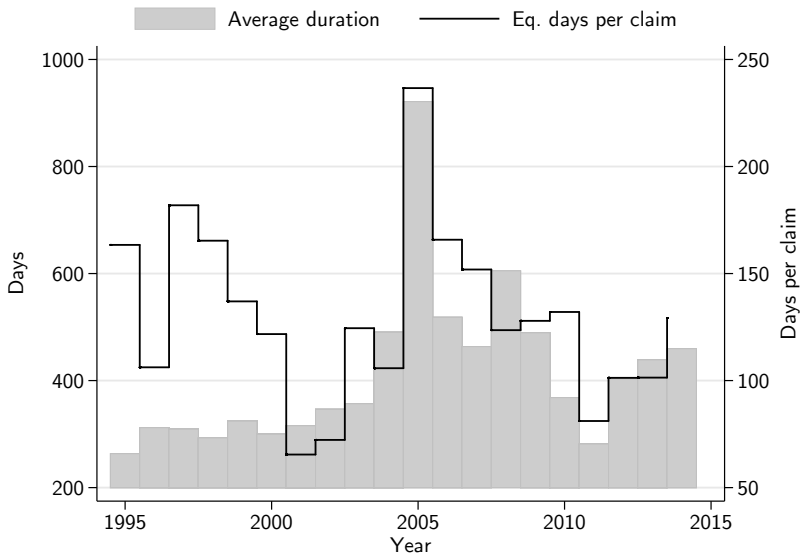
Figure 2: Length of panel proceedings and number of claims.



200 days, on average, and the the average time spent on panel reports increased steadily over the subsequent years. The total number of claims seem to also have weakly increased, but there is larger variation in this variable compared to unique claims (also seen in Figure A.4). The co-variation between duration and claim is moderate ($\rho = 0.3$).

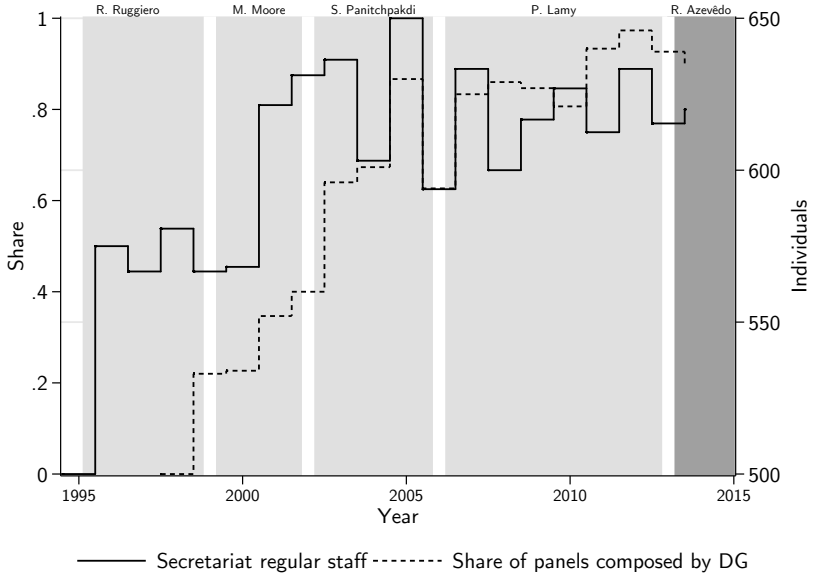
Figure 3 presents the two outcomes over time. The average duration for the first five years was rather low at around 302 days, while average Eq. days per claim was high, at 157, which can be explained by a low average number of claims of seven (see Figure 2).

Figure 3: Duration and Eq. days per claim, over time.



Although the number of requests for panel adjudication has been steady in the last years, Figure 4 shows that the Secretariat's panel composition duties have increased over time. If we take Figure 4 as evidence of a general increase in workload, we would expect that the time to complete panel proceedings would also have increased. But this is not reflected in Figure 2, suggesting, perhaps, that the increase in staff capacity has somewhat compensated the increasing demands of the DSM.

Figure 4: Staff capacity and panel composition



5 Results

The results in Table 3 reveal a combination of expected and unexpected effects. We will mainly interpret the coefficients as estimated statistical relationships since the non-significant effects are, in practice small. The one exception is Art. 8.10-panels, which is associated with a 22-days shorter duration on average. The main significant result is that former co-panelists can reduce duration by 68 days on average. The result holds at the median but the effect is smaller (37 days). This difference is driven by the two Airbus-Boeing disputes, and removing these in the OLS estimation results, instead, in a a 40-day reduction.[11] The small and positive effect of case law that we observe is also sensitive to the inclusion of Airbus-Boeing disputes, and if we remove these cases, the result is similar to the effect at the median. Panel experience does not appear to expedite the time required to

[11] Note that in-crowd is trivially correlated with panel experience. However, the latter has no impact on the in-crowd coefficient.

issue panel reports, which confirms the findings in Horn (2013) and Busch and Pelc (2009).

As expected, working on multiple panels concurrently increases duration by over 100 days (median). But it is unlikely that this would be the case for a permanent panel that handles multiple panels. For example, the AB consistently met their 90-day deadline up until 2014, when they were overwhelmed by a sudden increase in appeals. Appointed *ad hoc* panelists are hardly expected to work full-time on a case (Johannesson and Mavroidis, 2015), even if they are assigned multiple ones. Thus, it is difficult to calculate the trade-off in time between a full-time, and a more efficient, permanent panel and the increased workload it would entail.

Next, we discuss if experience and in-crowd are positively associated with efficiency. Unexpectedly, Eq. days per day appears to suffer with panel experience, though not by much. Each prior panel experience adds around 5 more Eq. days per claim. For 10 claims, that would amount to 16 extra days, although the effect is diminishing after a combined panel experience of twelve.[12] These effects are most likely explained by selection effects: As a panelist amasses experience, the chances of being appointed increase, but at the same time, panelists with experience are more likely to be assigned more complex and therefore time-consuming cases. The magnitudes are similar at the median, though not significant. We do not observe any positive effects on efficiency when working with previous co-panelists.

Lastly, panels that were appointed by the DG are associated with both a 40-day increase in duration and a 20-day increase in Eq. days per claim, but these results are not robust as they disappear at the median. However, when we remove the Airbus-Boeing disputes in OLS, the associated increase is instead 27 Eq. days per claim, but it is not statistically significant. Considering the steady increase in disputes that refer the panel composition to the DG, as seen in Figure 4, the relationship between DG-appointed panels and complexity may not be as strong as theorized. It is possible that the parties seek the assistance of the DG more often simply out of convenience and time considerations, rather than an increase in politically sensitive disputes.

Before concluding, it is important to keep in mind that we implicitly assumed constant quality when we interpreted the results. That is, efficiency gains are not achieved at the expense of quality in rulings. But, clearly, for some minimum value of Eq. days per claim,

[12] $\frac{\partial y^{x_d}}{x}\left[\gamma_1 x + \gamma_2 x^2\right] = 4.7 + 2 \cdot (-0.20x) = 0 \Leftrightarrow x = 12$

additional time-saving efforts will probably affect the quality of the panel report negatively.

Table 3: OLS regression

	Duration	Eq. days per claim
No. of claims	2.3*	−1.6***
	(1.0)	(0.25)
Expert	250***	93***
	(54)	(23)
Total Panel Experience	7.9	4.7'
	(7.5)	(2.6)
Total Panel Experience (squared)	−0.32	−0.20'
	(0.26)	(0.10)
Domestic instrument disputes	1.3	−3.0
	(24)	(11)
Case law	1.5*	0.13
	(0.71)	(0.25)
Case law || DI cases	0.34	−0.29
	(1.4)	(0.42)
No. of Third Parties	4.0'	−1.4'
	(2.3)	(0.75)
Appointed by DG	42'	21'
	(25)	(11)
Multiple panels	178*	49*
	(89)	(19)
In-crowd	−68*	−4.1
	(31)	(10)
In-crowd || DG	14	−5.3
	(15)	(8.9)
Art. 8.10 composition	−22	11
	(22)	(10)
(constant)	226***	126***
	(22)	(10)
R^2	0.36	0.41
N	210	210

*** $p < 0.1\%$, ** $p < 1\%$, * $p < 5\%$, ' $p < 10\%$

Table 4: Median regression (robustness)

	Duration	Eq. days per claim
No. of claims	2.0***	−1.4***
	(0.42)	(0.23)
Expert	215***	77***
	(35)	(20)
Total Panel Experience	5.6	3.7
	(5.2)	(2.9)
Total Panel Experience (squared)	−0.27	−0.18
	(0.28)	(0.16)
Domestic instrument disputes	6.0	6.9
	(23)	(13)
Case law	0.51	0.067
	(0.62)	(0.34)
Case law || DI cases	0.91	0.26
	(1.2)	(0.65)
No. of Third Parties	6.7***	−0.57
	(2.0)	(1.1)
Appointed by DG	2.1	5.8
	(23)	(13)
Multiple panels	105**	26
	(35)	(20)
In-crowd	−37*	3.4
	(19)	(10)
In-crowd || DG	12	−3.9
	(16)	(9.1)
Art. 8.10 composition	2.7	14
	(26)	(14)
(constant)	221***	106***
	(26)	(14)
N	210	210

*** $p < 0.1\%$, ** $p < 1\%$, * $p < 5\%$, ' $p < 10\%$

5.1 Robustness

Table 5: Comparison of coefficients

	(I)		(II)		(III)	
	OLS	Median	OLS	Median	OLS	Median
Duration	7.9	5.6	-68.4	-37.1	1.5	0.5
	(7.5)	(5.2)	(31.0)	(18.5)	(0.7)	(0.6)
Eq. days per claim	4.7	3.7	-4.1	3.4	0.1	0.1
	(2.6)	(2.9)	(10.2)	(10.4)	(0.2)	(0.3)
Days per claim (equal weight)	3.1	0.5	12.4	10.1	-0.1	-0.1
	(2.0)	(2.1)	(11.3)	(7.7)	(0.2)	(0.3)
Days per unique claim (equal weight)	2.5	0.9	11.9	11.4	0.4	0.2
	(2.8)	(3.7)	(12.9)	(13.4)	(0.3)	(0.4)
Eq. days per unique claim	3.6	3.3	-8.2	-9.3	0.8	0.4
	(3.6)	(4.0)	(14.4)	(14.5)	(0.3)	(0.5)

(I) Panel Experience (II) In-crowd (III) Case Law

We compare three panel-related variables: panel experience, In-crowd and case law across different variants of duration and efficiency (see Appendix E for full regression tables and Figure A.3 in Appendix A for variation over time). Then OLS coefficients remain insignificant for both panel experience and case law for all specifications. In-crowd, however, is sensitive to the functional form. If we apply equal weights to the claims, the variable in-crowd is associated with an increase of 12 days per claim, on average.

6 Discussion

The problem of delays in panel deliberations have been a long-standing issue, going as far back as the GATT (Tuthill et al., 1985). But in recent years, the problem has intensified with the success of the DSM. One suggestion on how to improve the system has been to replace *ad hoc* panels with permanent ones. This study offers some new insights on the association between time, efficiency and panel features that characterize permanent panels. Apart from using the time it takes for panels to issue panel reports to measure timeliness, we also measure efficiency, using Eq. days per claim, which is constructed by taking the

duration and divide it by the square root of number of claims. In this manner, we are able to account for economies of scale in examining claims.

We will focus our discussion on three main results. First, one of the main arguments for permanent panels is that experience matters for both the quality of panel reports and efficiency of panel work (Busch and Pelc, 2009), and permanent panels will ensure that judges accumulate such experience. Our results show, however, that panel experience does not have any substantial impact on duration. Although we found positive effects on Eq. days per claim—most likely due to selection effects—the effect is small, making it difficult to infer anything credible about its role in time-saving.

Second, as found previously in Horn (2013), case law has trivial effects on timeliness and efficiency. This is a rather counter-intuitive result. One explanation may be that there is no rule of either vertical or horizontal *stare decisis* in the DSM. That is, neither previous panel rulings nor the rulings of the Appellate Body are binding, although the panel is permitted to consult them. Hence, a large body of case law is not as helpful as one might think. There has been a long-standing belief that a *de facto stare decisis* has developed in the DSM, particularly since the AB has insisted that the panel (and the AB itself) follow AB jurisprudence (see DS 344, DS 322). Although there have been relatively few instances of explicit tension between the panel and AB rulings, aside from the mentioned disputes, the token effect of case law on timeliness may suggest that panels do not, in general, apply *stare decisis*. If this is the case, a panelist's experience should be even more important as case law could have been a way to support less experienced panelists.

Third, panels that consist of former co-panelists were associated with the largest reduction in duration. Two panelists that have worked together previously is associated with a 40-day decrease in the time it takes to issue panel reports. We do not have a causal framework, so there could be several explanations for this result apart from the straightforward one, that an established work relationship can expedite panel deliberations. We previously discussed that the lack of results from panel experience could be due to selection effects, and the same could be true for in-crowd; former co-panelists are more likely to be appointed specific disputes that, for various reasons, take a shorter amount of time to complete. However, it is difficult to surmise what type of dispute characteristics would cause such a mechanism.

References

Bourgeois, J. H. (2003). Comment on a WTO permanent panel body. *Journal of International Economic Law 6*(1), 211.

Busch, M. L. and K. J. Pelc (2009). Does the WTO need a permanent body of panelists? *Journal of International Economic Law 12*(3), 579.

Cottier, T. (2003). The WTO permanent panel body: a bridge too far? *Journal of International Economic Law 6*(1), 187.

Davey, W. J. (2003). The case for a WTO permanent panel body. *Journal of International Economic Law 6*(1), 177.

Horn, H. (2013). The time WTO panels require to issue reports. Working paper no. 979, Research Institute of Industrial Economics.

Horn, H., G. Maggi, and R. W. Staiger (2010). Trade agreements as endogenously incomplete contracts. *The American Economic Review 100*(1), 394–419.

Howse, R. (2012). Regulatory Measures. In A. Narlikar, M. Daunton, and R. M. Stern (Eds.), *The Oxford Handbook on the World Trade Organization.* Oxford University Press.

Hudec, R. E. (1999). The new WTO dispute settlement procedure: an overview of the first three years. *Minn. J. Global Trade 8*, 1.

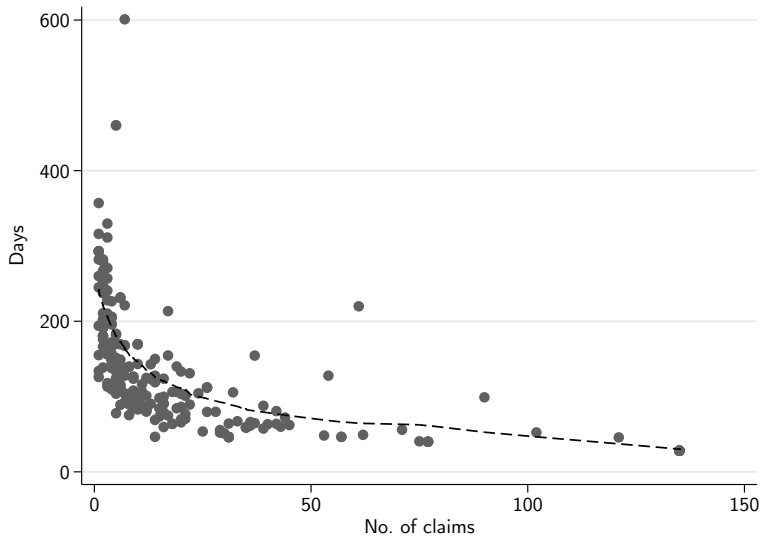
Jackson, J. H. (1979). Governmental disputes in international trade relations: a proposal in the context of gatt. *J. World Trade L. 13*, 1.

Johannesson, L. and P. C. Mavroidis (2015). Black Cat , White Cat : The Identity of the WTO Judges. *Journal of World Trade 49*(4), 685–698.

Tuthill, L. L., J. E. Guth, K. A. Skidmore, and P. Gibson (1985). *Review of the effectiveness of trade dispute settlement under the Gatt and the Tokyo Round agreements : report to the Committee on Finance, U.S. Senate, on investigation no. 332-212 under section 332(g) of the Tariff Act of 1930.* Number 1793 in USITC publication. Washington, DC : U.S. International Trade Commission.

A Data

Figure A.1: Square root equivalence scale



A.1 Variables

Figure A.4 plots the distribution of some of the variables that we include in the analysis. Note that the values are plotted against each dispute's DS number, which can be interpreted as time since these numbers are cumulative.

Figure A.2: Duration

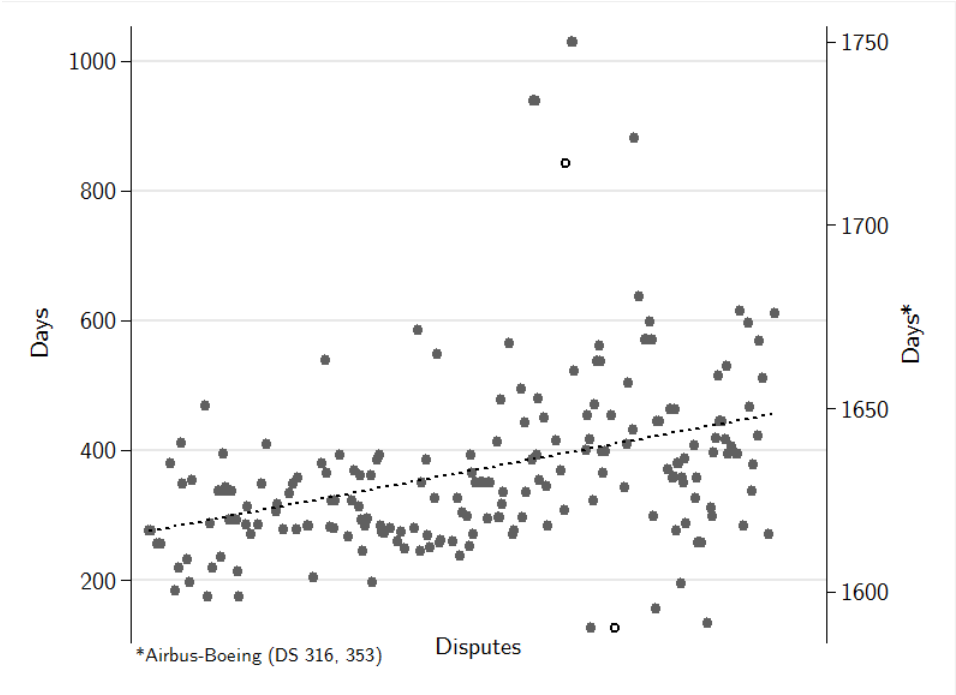


Figure A.3: Alternative outcomes

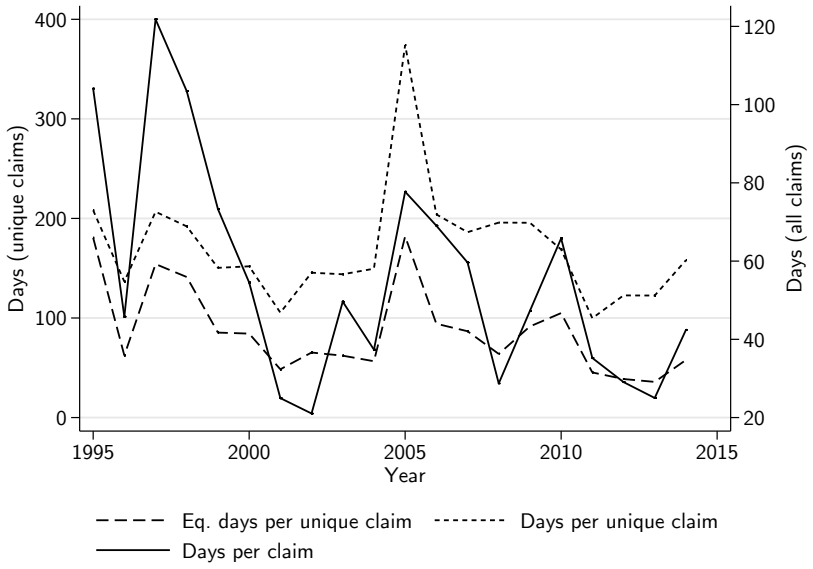
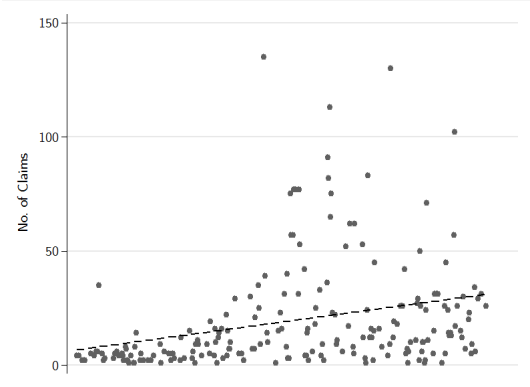
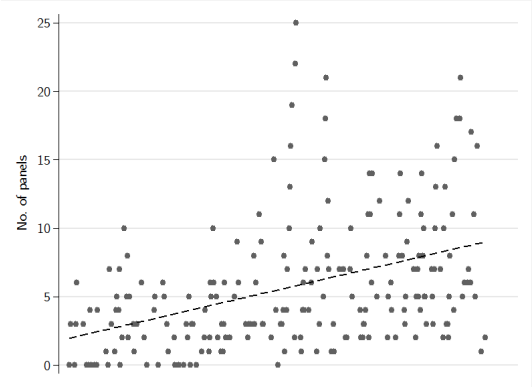


Figure A.4: Variables

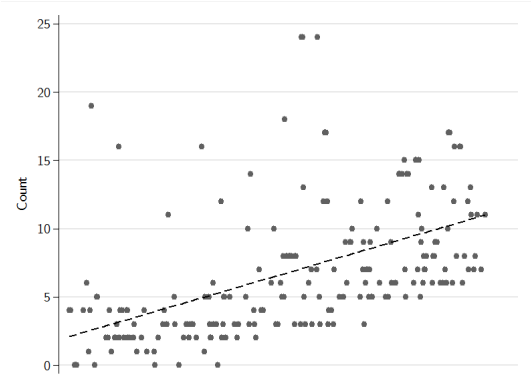
(a) Total number of claims



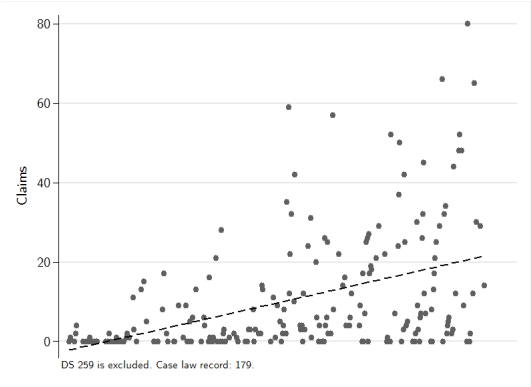
(b) Panel Experience



(c) Third parties

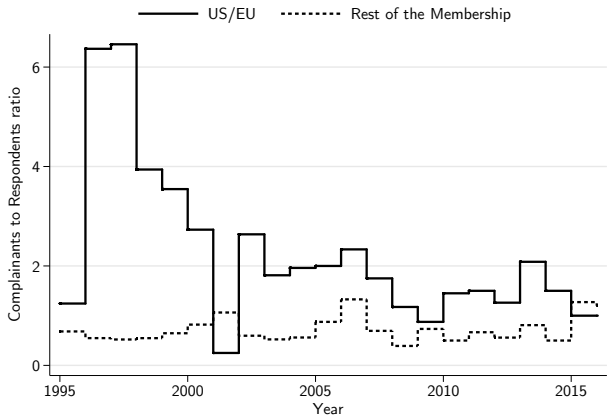


(d) Case Law



B Participation

Figure A.5: Request for consultations, complainants to respondents ratio



C Director-Generals

Table A.1 presents all WTO DGs and their respective terms (four years). The current DG, Roberto Azevêdo, was appointed in late 2013.[13]

Table A.1: Director-Generals

Director-General	Years	Nationality	#Panels	#Panels by DG
Renato Ruggiero	1995--1999	Italy	70	29
Mike Moore	2000--2002	New Zealand	48	36
Supachai Panitchpakdi	2003--2005	Thailand	33	27
Pascal Lamy	2006--2013	France	77	57
Roberto Azevêdo	2014--	Brazil	5	4

[13]The procedural guidelines to appoint a DG was adopted in 2002 and is described in WT/L/509.

D Country classification

Table A.2: Country Classification

IND	DEV
Australia	Antigua and Barbuda
Canada	Argentina
Japan	Chile
Korea	Chinese Taipei
Mexico	Colombia
New Zealand	Costa Rica
Norway	Dominican Republic
Poland	Ecuador
Switzerland	Egypt
Turkey	El Salvador
	Guatemala
G2	Honduras
EU	Indonesia
US	Malaysia
	Pakistan
BRIC	Panama
Brazil	Peru
China	Philippines
India	Thailand
	Ukraine
	Venezuela
	Viet Nam

E Alternative specifications

The outcomes in Table A.3 and A.4 are defined as follows: (1) days per claim, equal weights; (2) days per unique claim, equal weights and; (3) Eq. days per unique claim.

Table A.3: Alternative outcomes: OLS regression

	(1)	(2)	(3)
No. of claims	−1.3***	−1.1***	−0.83*
	(0.21)	(0.24)	(0.36)
Expert	39**	48**	111***
	(13)	(18)	(24)
Total Panel Experience	3.1	2.5	3.6
	(2.0)	(2.8)	(3.6)
Total Panel Experience (squared)	−0.13	−0.19	−0.24
	(0.093)	(0.13)	(0.15)
Domestic instrument disputes	−6.3	−8.4	−7.9
	(9.8)	(12)	(13)
Case law	−0.14	0.43	0.80*
	(0.19)	(0.26)	(0.35)
Case law || DI cases	−0.26	0.12	−0.026
	(0.34)	(0.64)	(0.68)
No. of Third Parties	−2.1**	−2.8**	−1.3
	(0.73)	(1.0)	(1.1)
Appointed by DG	10	22’	28*
	(11)	(12)	(13)
Multiple panels	19	42*	72*
	(14)	(21)	(34)
In-crowd	12	12	−8.2
	(11)	(13)	(14)
In-crowd || DG	−14	−13	−0.98
	(11)	(12)	(10)
Art. 8.10 composition	18’	15	−0.11
	(11)	(17)	(14)
(constant)	72***	85***	133***
	(11)	(13)	(12)
R^2	0.35	0.25	0.23
N	210	210	210

*** $p < 0.1\%$, ** $p < 1\%$, * $p < 5\%$, ’ $p < 10\%$

Table A.4: Alternative outcomes: median regression

	(1)	(2)	(3)
No. of claims	−0.81***	−0.74*	−0.81*
	(0.17)	(0.30)	(0.33)
Expert	32*	68**	152***
	(15)	(25)	(27)
Total Panel Experience	0.52	0.95	3.3
	(2.1)	(3.7)	(4.0)
Total Panel Experience (squared)	−0.013	−0.086	−0.22
	(0.12)	(0.20)	(0.22)
Domestic instrument disputes	4.4	−4.4	−12
	(9.4)	(16)	(18)
Case law	−0.13	0.19	0.44
	(0.26)	(0.44)	(0.48)
Case law || DI cases	−0.20	−0.20	0.60
	(0.49)	(0.84)	(0.92)
No. of Third Parties	−1.2	−0.40	−0.63
	(0.83)	(1.4)	(1.6)
Appointed by DG	5.9	6.2	−1.2
	(9.7)	(17)	(18)
Multiple panels	3.2	−2.5	37
	(15)	(25)	(27)
In-crowd	10	11	−9.3
	(7.7)	(13)	(15)
In-crowd || DG	−8.0	−2.9	14
	(6.8)	(12)	(13)
Art. 8.10 composition	8.1	−1.8	−6.8
	(11)	(18)	(20)
(constant)	47***	54**	128***
	(11)	(19)	(20)
N	210	210	210

*** $p < 0.1\%$, ** $p < 1\%$, * $p < 5\%$, ' $p < 10\%$