Backtracking Under Scrutiny: Evidence and Regulation in the WTO

Shing-hon Lam*

1 January 2026

Abstract

Regulatory trade barriers have proliferated, yet formal WTO dispute settlement is rarely used to contest them. How do governments discipline trade-restrictive regulations that claim legitimate objectives? I argue that the WTO constrains regulatory protectionism primarily through transparency and peer scrutiny rather than adjudication. In the notification process, governments choose how much justificatory evidence to disclose, and other members infer motives from this disclosure. Submitting little or no evidence makes protectionist intent salient and increases the likelihood of comments and early modification. Submitting extensive documentation can also backfire: under scientific uncertainty, large and heterogeneous evidence bundles may appear overinclusive or strategically constructed, creating additional points of contestation and raising doubts about sincerity. As a result, regulatory credibility follows a nonlinear pattern in which intermediate levels of disclosure attract the least attention, while both minimal and excessive disclosure heighten scrutiny and the likelihood of revision, especially when protectionist incentives are strong. I test these claims using an original dataset of 50,599 WTO TBT and SPS notifications from 2010–2024 and a new dataset of regulatory comments from the European Union's TBT comment portal. Consistent with the theory, governments are more likely to modify proposed regulations when they submit low or high evidence and when protectionist pressure is stronger; medium-evidence measures receive the fewest comments; receiving a comment substantially increases the likelihood of modification; and higher evidentiary volume reduces comment incidence for protectionist measures but, conditional on receiving a comment, makes high-evidence protectionist measures especially likely to be revised. These findings show how multilateral oversight can discipline regulatory protectionism even when legal enforcement is weakened.

*PhD candidate at the Department of Political Science, UCLA. shinghon@ucla.edu.

1 Introduction

In 2016, new French compostability requirements caused biosourced plastic bags shipped from Asia to decompose in transit, raising concerns of disguised protectionism. Because roughly 90 percent of such bags had previously been imported from Asia, the measure appeared to restrict market access without overtly violating trade rules (de Tannenberg, 2016). How can trading partners such as China contest trade-restrictive regulations that claim legitimate objectives?[1] Traditionally, governments contested regulatory trade barriers through negotiation, WTO consultations, and ultimately dispute settlement. Yet this pathway has become increasingly constrained. WTO case law now permits differential treatment of imported and domestic products when grounded in a genuine regulatory objective, narrowing the scope for successful legal challenges (Howse and Langille, 2023). At the same time, rising legal complexity has made dispute settlement costly and inaccessible for many members, particularly developing countries (Pauwelyn and Zhang, 2018). These limits are compounded by the paralysis of the WTO Appellate Body. As a result, legal avenues for contesting suspect regulations are often unavailable.

Conventional accounts of the global trade regime rest on a clear ideal-world logic: cooperation requires credible enforcement. The transition from the GATT to the WTO is therefore understood as a response to the limits of diplomacy, strengthening dispute settlement and authorized retaliation to deter protectionism (Rosendorff and Milner, 2001; Rosendorff, 2005; Pelc, 2016). Even negotiated outcomes are understood to occur in the shadow of dispute settlement consultations and potential panel rulings (Busch, 2000; Busch and Reinhardt, 2000). From this perspective, the erosion of enforcement has led many observers to conclude that the WTO's rule-based system is increasingly ineffective. Yet regulatory politics tell a different story. Under the TBT and SPS agreements, only 113 dispute settlement consultations have been requested, and just 24 cases have reached a ruling. Nevertheless, using data on 50,599 Technical Barriers to Trade (TBT) and Sanitary and Phytosanitary (SPS) notifications from 2010–2024, I show that governments modified 3,434 proposed regulations—about 6 percent of all cases (Figure 1). This pattern coincides with extensive peer scrutiny through transparency mechanisms: since 2010, the European Union has submitted formal comments on 1,431 foreign regulations, while other countries have commented on 326 EU measures. Because

[1] Regulatory trade barriers of this kind have become increasingly common. (see also Meyer, 2022). Under the European Union's proposed deforestation law, exporting a single 20-foot container of palm oil would require an estimated 1.2 million documents, prompting major agricultural exporters to denounce the measure as protectionist (Hancock and Ruehl, 2023; Ruehl, Hancock, and Terazono, 2023). In the United States, new energy-efficiency standards will require household dishwashers sold after 2027 to reduce energy use by 27 percent and water use by 34 percent in standard cycles—changes expected to raise production costs significantly and already questioned by trading partners (Lange, 2023).

regulatory standards now govern market access across core policy domains, explaining why such modifications occur is central to understanding how trade institutions continue to constrain protectionism—thereby protecting consumer welfare—despite weakened legal enforcement.

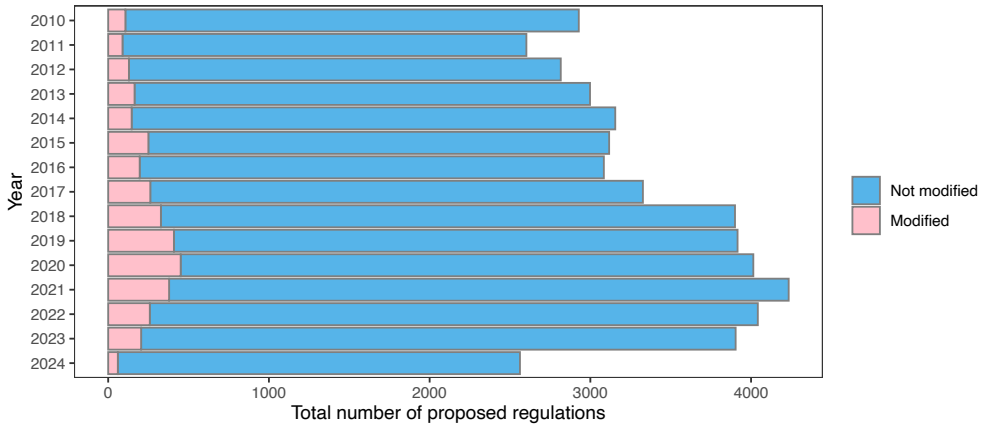


Figure 1: Total number of proposed WTO TBT/SPS regulations by year (2010–2024).

In this paper, I ask why governments modify or withdraw regulatory trade barriers in response to international scrutiny. I argue that the WTO functions as an epistemic regime[2] in which governments signal the legitimacy of regulatory measures through the volume of evidence submitted during the notification process. Because regulatory intent is unobservable, other members infer motives from justificatory disclosure. Submitting little or no evidence makes protectionist intent salient, increasing the likelihood of early modification. Submitting extensive documentation can also provoke scrutiny: under scientific uncertainty, regulators often cite numerous studies, standards, and expert reports, yet to outsiders these dense and heterogeneous evidence bundles may appear strategic or overcompensatory, raising doubts about sincerity. Regulatory credibility therefore follows a non-linear pattern, with intermediate levels of evidence attracting the least attention. Regulations supported by either minimal or excessive evidence are more likely to be modified—especially when protectionist incentives are present—because both extremes invite scrutiny and undermine perceived legitimacy.

This signaling logic operates through peer review rather than adjudication. Other members may respond to notifications by submitting comments, and both very low and very high levels of evidence are more likely to attract such scrutiny than intermediate disclosure, which appears least suspicious. Comments matter because they impose tangible costs even in the absence of formal enforcement. They signal potential

[2] See Meyer (2021, 55–56) for characterising international organisations as an epistemic regime.

retaliation, raise reputational concerns about regulatory bad faith, and threaten valuable trade relationships, making regulatory revision more likely. Anticipating these pressures, protectionist governments may inflate evidentiary submissions ex ante to deter comments by raising the cost of review or obscuring intent through technical complexity. Once a comment is filed, however, incentives reverse. Extensive documentation supplies challengers with more points of contestation, increases defense costs, and heightens the risk of reversal, making high-evidence protectionist measures especially likely to be modified.

To evaluate this argument, I assemble an original dataset of 50,599 notifications submitted under the WTO's Technical Barriers to Trade (TBT) and Sanitary and Phytosanitary (SPS) agreements between 2010 and 2024. Each observation corresponds to a proposed regulation and tracks whether it was subsequently modified or withdrawn based on follow-up addenda and corrigenda. The key explanatory variable is the volume of justificatory evidence submitted with each notification, proxied by the number of pages attached, which I combine with product-level trade data to capture protectionist incentives. To measure international scrutiny short of formal dispute settlement, I further construct an original dataset of regulatory comments from the European Union's TBT comment portal, which records formal comments submitted by the EU on other members' notifications as well as comments filed by foreign governments on EU regulations since 2010. The empirical results align with the theory's predictions, showing that both minimal and excessive evidentiary disclosure are associated with higher rates of regulatory modification—especially under conditions of protectionist pressure—and that peer comments play a central mediating role.

This paper contributes to broader debates on how international organizations influence state behavior in the absence of effective legal enforcement. Existing research emphasizes peer pressure, monitoring, and reputational incentives as mechanisms through which IOs shape compliance (e.g. Pevehouse, 2005; Kelley, 2004, 2007; Simmons, 2000; Tomz, 2007). I extend this literature by showing that transparency is not merely a procedural requirement but a strategic instrument: the volume and character of evidentiary disclosure shape how regulations are interpreted, scrutinized, and revised. While international law scholars highlight the importance of scientific justification in assessing discrimination (e.g. Rigod, 2015; Lester, 2022), this paper demonstrates that how evidence is presented within institutionalized review processes is itself consequential. By theorizing the WTO as an epistemic regime, the findings show how regulatory discipline can emerge through peer review and reputational dynamics rather than formal adjudication, underscoring a broader role for international organizations in structuring justification and accountability.

The remainder of the paper proceeds as follows. Section 2 outlines the WTO notification and comment

process for regulatory measures. Section 3 presents the core theory, explaining how evidentiary disclosure signals regulatory legitimacy. Section 4 unpacks the underlying mechanisms, focusing on how peer review and comments translate disclosure into regulatory revision. Section 5 describes the original datasets of over 50,000 notifications under the TBT/SPS agreements and the EU TBT comment portal. Section 6 presents the empirical results, and Section 7 concludes.

2 WTO's three pre-settlement stages

Formal dispute settlement represents only the final stage of WTO scrutiny. In practice, regulatory measures are reviewed through three pre-settlement stages—notification, commenting, and committee-level discussion—that often shape outcomes before litigation becomes relevant. The questioning, objection, and information exchange in these stages can induce governments to revise or withdraw contested regulations.

The first stage is notification. Before implementing any regulation that may have a significant effect on trade, members are required to notify other governments of the measure's objectives, expected trade impact, and scope of affected products and partners (WTO, 2018). Notifications can attach documents of evidence such as scientific studies, laboratory tests, inspection protocols, or risk-assessment models. TBT notifications are closely linked to subsequent specific trade concerns, indicating that notification itself triggers scrutiny (Horn, Mavroidis, and Wijkström, 2013). The second stage is commenting. Other members may submit comments requesting clarification, raising objections, or proposing amendments to the notified measure. Regulators are required to respond to these comments and are encouraged to publish their responses before finalizing the regulation (WTO, 2018). Although comments carry no formal legal force, they often trigger substantive engagement, including the provision of draft legal text, scientific opinions, or commitments to regulatory flexibility. Through this exchange, the commenting process embeds expectations of transparency and responsiveness into domestic rulemaking (Downes, 2012; Karttunen, 2020).

The third stage involves Specific Trade Concerns (STCs) raised in the WTO's TBT and SPS Committees. Because STCs face few procedural barriers, they are widely used by governments seeking dialogue without escalating to formal disputes and to attract potential third-party involvement (Busch and Pelc, 2015; Manak, 2019). Committee discussions are typically deliberative rather than adversarial: officials exchange technical information, clarify regulatory intent, and debate scientific risk assessments. Many STCs are resolved once sufficient clarification or technical justification is provided, or when regulators agree to amend aspects of the measure (Lang and Scott, 2009). Even without legal compulsion, the reputational costs of committee-level

scrutiny and the professional norms surrounding scientific risk assessment often pressure governments to align regulations more closely with WTO expectations (Downes, 2012; Cho, 2011; Wolfe, 2020).

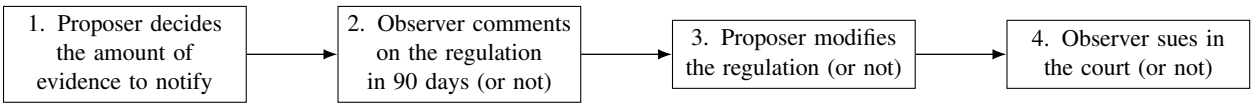


Figure 2: Sequence of the WTO's three pre-settlement stages

I conceptualize the WTO's three pre-settlement stages as a sequence. A government first notifies a proposed regulation and decides how much supporting evidence to submit. Other members then have a fixed period—typically 90 days—to review the notification and decide whether to comment. In response to any comments received, the notifying government may modify the proposed regulation or proceed unchanged, with formal dispute settlement remaining a final and infrequently used option.

This sequence is accompanied by an ongoing discourse of scientific evidence, in which regulatory legitimacy is articulated, contested, and revised through text. First, in preparing regulations, policymakers frame regulatory design as conditional on scientific assessment conducted prior to decision-making. As EU health commissioner Olivér Várhelyi stated, "if science says it is not safe, then we shouldn't have it. If we want to be scientifically based, then science is universal." Officials further emphasized that proposed restrictions target "only the most hazardous pesticides" and that "there will be an impact assessment to protect competitiveness before any decisions" (Bounds, 2025). Second, at the notification stage, evidence is formally embedded in regulatory infrastructure. For example, EU notification G/TBT/N/EU/1098 lists as "Relevant documents: Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH Regulation)," alongside the "Registry of restriction intentions until outcome – ECHA."

Third, during the commenting phase, other members frame scrutiny as an evidentiary demand rather than a legal accusation. In a written comment on a U.S. notification, China "appreciates the USA for fulfilling the transparency obligation under WTO, as well as for the opportunities for other WTO Members to make comments," before recommending that "the United States conduct further research on the degrees of harms caused by non-mercury metal hazardous air pollutants… further improving the setting of relevant limits" (China WTO/TBT National Notification & Enquiry Center, 2022). Finally, in response to such scrutiny, governments explicitly link regulatory persistence or relaxation to ongoing evidence production. Hong Kong stated that it "will review whether the current measures can be gradually relaxed according to

the scientific evidence available," and noting that "two independent samplings were carried out in October 2024 and February 2025 under the monitoring arrangement within the IAEA framework" (WTO Trade Concerns Database, 2025). Taken together, these exchanges show that scientific evidence circulate continuously throughout the pre-settlement process, shaping how regulations are interpreted, challenged, and revised.

3 Evidence and regulatory credibility in the WTO

The preceding section shows that regulatory contestation in the WTO unfolds through a structured sequence in which scientific evidence is assembled, evaluated, and revised under international scrutiny. These institutional features create incentives for governments to deploy evidentiary disclosure strategically. Because regulators' underlying motives—whether legitimate or protectionist—are not directly observable, other members infer intent from the amount and character of evidence submitted. The theory developed in this section formalizes this logic by modeling evidentiary disclosure as a signaling device through which governments manage perceptions of regulatory legitimacy, and by deriving how different levels of disclosure shape scrutiny and the likelihood of regulatory modification. The full formal model and proofs are presented in Appendix A2.

Scientific evidence occupies a central role in how regulatory trade barriers are presented and assessed in the WTO. The requirement that trade-restrictive measures rest on scientific justification is widely understood as a safeguard against protectionism (Herwig, 2008). Scientific reasoning legitimates regulation by constraining arbitrary decision-making and signaling credibility, while still allowing governments discretion over acceptable levels of risk (Howse, 2000). Yet science rarely yields determinate policy prescriptions. Risk assessment necessarily embeds judgments about uncertainty and acceptable harm and is shaped by institutional and political context rather than purely technical criteria (Walker, 2003; Busch et al., 2004). In WTO practice, what counts as relevant science is co-produced through legal procedures—panels select experts, structure inquiry, and interpret evidence—so evidentiary submissions are evaluated rather than taken at face value (Jasanoff, 2004, 2008). This indeterminacy creates scope for signaling: because regulatory intent is unobservable, other members rely on evidentiary disclosure to infer legitimacy.

When a government submits little or no scientific risk evidence, that absence itself becomes informative. Trade-restrictive measures lacking scientific justification are readily interpreted as disguised protectionism (Sykes, 2002), making regulatory intent legible to observers. As U.S. Trade Representative Robert Lighthizer put it, "making every regulation science-based is the equivalent of getting rid of protectionism"

(C-SPAN, 2020). Anticipating that weakly justified measures will attract scrutiny and perform poorly if challenged, regulators proposing low-evidence measures are therefore more likely to withdraw or modify them early in the process. Accordingly, countries are more likely to modify regulations when they submit low levels of scientific evidence, particularly when those regulations serve protectionist interests.

> H_{1a}: Countries are more likely to modify the regulation when they submit low evidence and (H_{1b}) also when they propose protectionist regulation.

Crucially, this logic does not imply that more evidence is always better. Under scientific uncertainty, regulators often face competing findings about risks or mitigation strategies, prompting them to submit large bundles of studies, expert opinions, and technical reports. What is intended as thorough or honest disclosure may instead appear to external observers as overloaded, unfocused, or strategically overinclusive, inviting suspicion rather than reassurance. Adjudication illustrates how impartial bodies evaluate such evidentiary volume. In the EC–Hormones dispute, the European Communities relied on numerous scientific sources—including IARC Monographs and "articles and opinions of individual scientists," which the Appellate Body acknowledged "do indeed show the existence of a general risk of cancer" (WTO Appellate Body, 1998, para. 200). Yet it emphasized that these materials "do not focus on and do not address the particular kind of risk here at stake," namely the carcinogenic or genotoxic potential of hormone residues in treated meat (Ibid.). The evidence was thus deemed "relevant but not sufficiently specific to the case at hand," and the Appellate Body concluded that "no risk assessment that reasonably supports or warrants" the measure had been provided (Ibid.). Rather than resolving uncertainty, the volume and diversity of scientific material failed to establish a coherent evidentiary narrative.

Regulatory authorities themselves also recognize that extensive documentation can provoke heightened scrutiny. Chinese officials report that other members "frequently questioned the technical details of our measures — particularly when the content was highly detailed or supported by extensive evidence materials," and that regulations with "detailed content and dense standards attracted excessive attention."[3] Concrete

[3] 2020 報告（上冊）— WTO/TBT-SPS 中國國家諮詢點報告（第 172 號 上冊）：「在受理、翻譯評議意見並處理答覆的過程中，我們感到其他成員對我方法規的技術細節提出頻繁質疑，尤其當法規內容較為詳盡或證據材料較多時。」 *We observed that other members frequently questioned the technical details of our measures—particularly when the content was highly detailed or supported by extensive evidence materials.* 2020 報告（下冊）— WTO/TBT-SPS 中國國家諮詢點報告（第 172 號 下冊）：「在受理過程中我們感到，一些法規因為內容詳盡、標準過密，引起了成員的過度關注。」 *During the review process, we found that some regulations, because of their detailed content and dense standards, attracted excessive attention from members.* 2020 年 TBT 評議匯編 — WTO/TBT-SPS 中國國家諮詢點報告（第 176 號）：「我們注意到，在對外評議中，部分成員對中國法規條款過於詳盡、數據支撐材料繁複表示懷疑，認為這可能構成貿易障礙。」 *We noticed that in external evaluations, some members expressed suspicion that China's overly detailed provisions and voluminous supporting data could constitute trade barriers.*

notification exchanges reveal the same dynamic. In response to China's notification G/TBT/N/CHN/1094, the European Union questioned why the draft relied on "three different test approaches," treating evidentiary multiplicity itself as a source of doubt. A similar pattern appears in reactions to EU notification G/TBT/N/EU/908. Although the EU supported its measure with extensive EFSA-based documentation, the United States dissected the evidence component by component, noting "several limitations in the risk assessments cited by the EU" and emphasizing that they covered only a "small number of pollinator species" under "limited crops, geographic, and climatic conditions." The reviewer ultimately reframed the scientific justification as provisional, concluding that the assessments were "based on very limited circumstances and with a significant number of uncertainties, as EFSA has acknowledged."

Once such elaborate evidentiary defenses unravel, maintaining the measure becomes costly. Extensive documentation publicly commits regulators to a detailed justification; when gaps or inconsistencies are exposed, credibility erodes and continued defense requires further costly clarification. Under these conditions, modifying or withdrawing the regulation becomes the least costly way to contain reputational damage and signal responsiveness. As with insufficient disclosure, excessive documentation can therefore signal contested intent, invite intensified scrutiny, and increase the likelihood of regulatory modification rather than insulating measures from challenge.

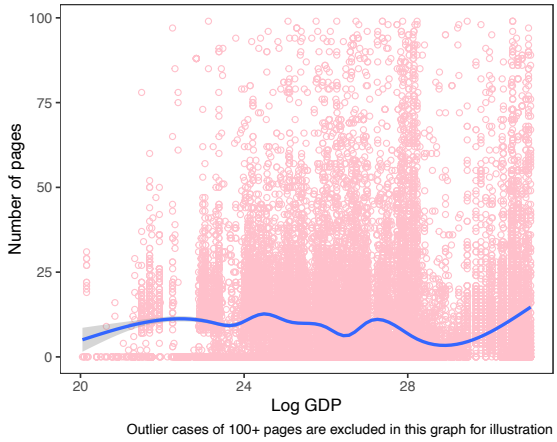
> H_{2a}: Countries are more likely to modify regulations when they submit excessively high levels of evidence, and (H_{2b}) especially when those regulations serve protectionist interests.

4 Peer scrutiny and regulatory modification

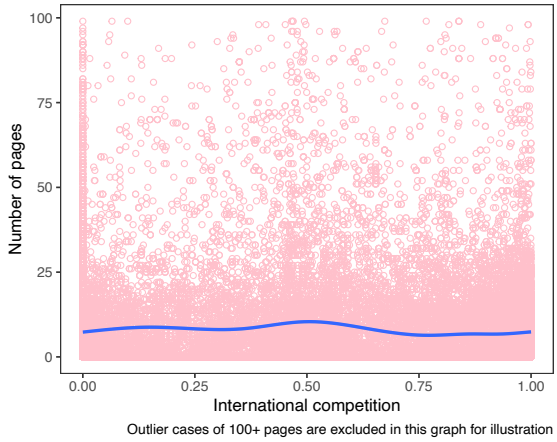
This section opens the black box linking evidentiary disclosure to regulatory modification by focusing on the role of peer comments. After a government notifies a regulation and its supporting evidence, other WTO members may respond through the comment process, creating an intermediate stage of scrutiny prior to any formal dispute. Because both very low and very high levels of evidence can appear suspicious, regulations supported by intermediate levels of documentation should attract the fewest comments. Comments themselves function as a key enforcement channel: once a regulation is publicly questioned, the proposing government faces reputational, relational, and strategic costs. Ignoring comments risks retaliation in other areas, signals bad-faith regulation, and can strain valuable trade relationships. As a result, regulations that receive comments are substantially more likely to be modified than those that do not, making peer review a central mechanism through which evidentiary disclosure translates into regulatory modification.

H_3: Medium-evidence regulations attract the fewest comments. H_4: Countries are more likely to modify regulations that received comments.

A natural question is why governments do not simply converge on intermediate levels of evidentiary disclosure, which attract the least scrutiny. Empirically, they do not. Figure 3a shows that evidentiary volume is not explained by state capacity: higher-income countries are no more likely than lower-income countries to submit moderate amounts of documentation, and variation in page length persists across income levels. Figure 3b similarly demonstrates that international competition does not account for disclosure choices: the degree of import competition faced by domestic producers is not systematically associated with how much evidence governments provide. These figures indicate that evidentiary disclosure is neither a mechanical function of resources nor of market exposure, but a strategic choice shaped by how governments anticipate scrutiny and manage regulatory risk.



(a) State capacity does not explain the amount of evidence provided



(b) International competition does not explain the amount of evidence provided

Figure 3: Determinants of evidentiary volume

Governments may sometimes choose to submit minimal or no scientific evidence as a deliberate signal. First, limited disclosure can indicate low resolve: producing detailed evidence is costly, and regulators may judge that the measure is not worth defending aggressively if challenged. In this sense, sparse justification signals a willingness to concede rather than escalate. Second, low evidence may reflect confidence that the regulation is obvious, routine, or unlikely to attract scrutiny, reducing the perceived need for extensive documentation. In both cases, minimal evidentiary submission communicates limited commitment to the measure and a readiness to modify or withdraw it if objections arise.

While regulations supported by intermediate levels of evidence attract the fewest comments, the incentives facing protectionist regulators differ. Too little evidence makes protectionist intent transparent, while too much evidence invites scrutiny. For protectionist measures, however, submitting unusually high volumes of documentation can be strategically attractive ex ante. By inflating evidentiary submissions, regulators raise the cost of review, increase technical complexity, and make it more difficult for other members to identify clear grounds for objection. Dense and highly technical dossiers can obscure intent and deter engagement, reducing the likelihood that comments are filed during the notification stage. As a result, increases in evidentiary volume are expected to reduce the probability that protectionist regulations receive comments relative to non-protectionist measures.

> H_5: As evidentiary volume increases, protectionist regulations become less likely to receive comments relative to non-protectionist regulations.

Once a comment is filed, the strategic value of extensive evidentiary disclosure reverses. High-evidence submissions provide challengers with more points of contestation, lock regulators into detailed justifications, and raise the costs of defense as inconsistencies or gaps are exposed. Rather than deterring scrutiny, large dossiers increase detection and reversal risk once review begins. At this stage, the comment process activates enforcement logics that operate even in the absence of formal adjudication: governments face reputational concerns, fear reciprocal retaliation, and seek to preserve ongoing trade relationships. For protectionist regulators, extensive documentation thus shifts from a tool of deterrence to a source of legal and political exposure. Under these conditions, modification or withdrawal becomes the least costly response, implying that conditional on receiving a comment, high-evidence protectionist measures are more likely to be modified than comparable low- or medium-evidence regulations.

> H_6: Once a regulation is commented on, high-evidence protectionist measures are more likely to be modified.

5 Data

To test the observable implications, I obtain 50,599 notifications from the online WTO-maintained TBT/SPS `epingalert.org` database between January 2010 and August 2024. I first classify the regulations using the document symbol — with or without "Add." or "Corr.". Members will initially submit the notifications with their initially proposed regulatory trade barriers, for example `G/TBT/N/UGA/680`. They may submit

subsequent notifications, for example `G/TBT/N/UGA/680/Add.1` to inform the status of the regulation. The statuses may include notifying the adoption, modification or withdrawal of the regulation, or they may provide additional information or amend the commenting period for the regulation. I read the documents with symbol "Add." or "Corr." to determine if they are notifying the adoption. My variable of interest is whether the member will adopt the initial regulatory trade barriers. I use each regulation as unit of analysis: if the regulation does not have subsequent addendum or corrigendum, or if the addendum only states that the regulation has been adopted,it is classified as not modifying the proposed regulation; otherwise, it is classified as modifying the proposed regulation.

My main explanatory variable is the amount of evidence presented by the proposing country. I proxy evidentiary disclosure using the number of pages in the documents that governments upload to the WTO alongside their notifications. Page length is a meaningful indicator because the types of materials relevant under the WTO's TBT and SPS agreements—scientific risk assessments, laboratory studies, epidemiological data, regulatory impact analyses, and legal justifications—are inherently document-intensive. Producing and compiling such materials requires administrative effort and technical capacity, making longer submissions a costly signal of evidentiary investment. As a result, page counts capture cross-country and within-country variation in the volume of scientific and legal material deployed to support a proposed regulation, thereby reflecting its evidentiary weight under WTO scrutiny.

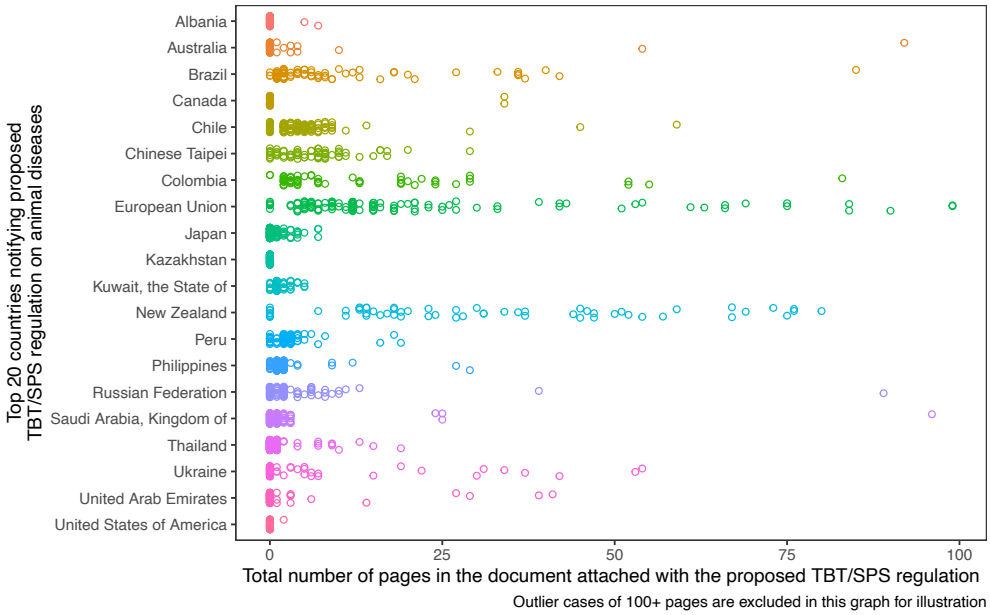


Figure 4: The scientific evidence in proposed TBT/SPS regulations on animal diseases

The amount of evidence varies substantially not only across issue areas but also across countries. For example, Figure 4 displays the top twenty countries making the most notifications on animal diseases. Brazil submitted eighteen pages of documentation for bird's eggs but only one page for turtle eggs, whereas Albania consistently files notifications on animal diseases without uploading any supporting documents. At the other extreme, the European Union submitted a total of 1,062 pages of documentation, reiterating the same health concerns cited in its own references on 753 occasions (G/SPS/N/EU/506). This wide dispersion in page length highlights the strategic nature of evidentiary disclosure and underscores its relevance as a measure of governments' signaling behavior in the WTO.

To infer whether a proposed regulation is more likely to be legitimate or protectionist, I exploit variation in international competition at the country–product level. I match each product listed in a WTO notification to BACI trade data and construct a measure of international competition as the ratio of imports to the sum of imports and exports. This measure captures the extent of foreign competitive pressure faced by domestic producers and is well suited to the task because BACI records all cross-border transactions, providing a comprehensive and comparable account of market exposure across products and countries.[4]

Low values of international competition (e.g., 0.1–0.3) typically reflect domestic self-sufficiency or limited exposure to foreign producers, while high values (e.g., 0.6–0.9) indicate openness and reliance on imported supply. In contrast, values near 0.5 characterize markets in which import-competing domestic firms and foreign producers coexist, creating the strongest incentives for selective trade restriction. Regulations affecting products in this middle range are therefore more likely to reflect protectionist intent, as they are consistent with limiting competitive pressure without fully closing the market. Importantly, as shown in Figure 3b, this inferred protectionist pressure does not mechanically translate into greater evidentiary volume: protectionist type has no discernible effect on the number of pages countries submit in their supporting documents, consistent with the model's emphasis on strategic evidentiary disclosure.

To explore how the complexity of regulatory documents interacts with protectionist motives, I estimate a regression using categorical variable for document length. The variable groups the pages into three bins: 0 for attached documents with zero pages (20,292 observations), 1 for those with 1–10 pages (13,763 ob-

[4] In the formal model, regulatory type—legitimate or protectionist—is private information and not directly observable by other WTO members. The empirical measure does not attempt to observe type itself. Instead, international competition serves as a proxy for underlying incentives: products with intermediate import penetration are those in which import-competing domestic firms and foreign producers coexist, generating the strongest political pressure for selective trade restriction. The measure therefore shifts the prior probability that a regulation is protectionist, while remaining consistent with the model's assumption that type is latent and inferred only imperfectly from observable signals.

servations), and 2 for those exceeding 10 pages (16,544 observations). These categories capture meaningful variation in the informational content of proposals (Figure 5).

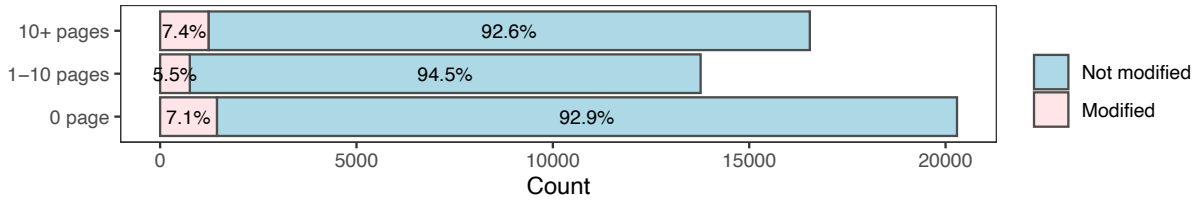


Figure 5: Cross tabulation of page numbers and regulation status

I include both a linear and a quadratic term for international competition to allow for a non-linear relationship between trade exposure and protectionist incentives, with the effect centered around intermediate values of import penetration (approximately 0.5 on the 0–1 scale), where protectionist pressures are hypothesized to be strongest. By interacting the evidence categories (`Medium` and `High`) with both the linear and quadratic terms of international competition, the model allows the effect of protectionist incentives on regulatory outcomes to vary across different levels of evidentiary disclosure. This specification directly tests whether moderate levels of evidence attenuate scrutiny while excessive evidentiary disclosure either fails to do so or reverses its effect. As robustness checks, I estimate a sequence of models that progressively introduce additional controls and fixed effects. These include a control for the country's log GDP to account for differences in administrative capacity and regulatory discretion, harmonized system (HS) chapter fixed effects to absorb unobserved heterogeneity across broad product categories, and year fixed effects to capture common shocks to WTO scrutiny and regulatory practices over time.

To examine how evidentiary disclosure translates into actual peer scrutiny within the WTO, I then draw on comment data from the EU Technical Barriers to Trade comment portal. The EU is the only WTO member that systematically publishes both the comments it submits on other members' notifications and the comments it receives on its own notifications, which allows for consistent and transparent comparison across countries. The dataset covers the period since 2010 and includes EU comments on other members' notifications as well as comments submitted by other WTO members on EU notifications. I code a notification as receiving a recorded comment if any comment appears in the portal. Overall, 5.5 percent of notifications in the sample receive at least one recorded comment.

6 Results

Consistent with H_{1a} and H_{2a}, the probability that a regulation is modified increases sharply when protectionist pressure coincides with either low or high evidentiary disclosure. As shown in Figure 6, moving from low international competition (approximately 0) to intermediate competition (around 0.5) is associated with an increase of roughly 6 percentage points—about a 150 percent rise—in the likelihood of modification. In addition, holding protectionist pressure constant, moving from medium to high levels of evidentiary disclosure increases the probability of modification by about 3 percentage points, or roughly 40 percent.

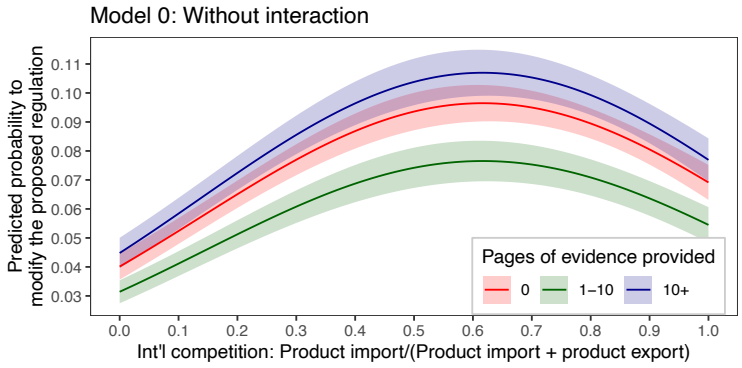


Figure 6: Predicted probability of regulatory modification by international competition and evidentiary volume (Model 0: without interation)

The results in Table 2 in the Appendix indicate a non-monotonic relationship between import penetration and the predicted probability of regulatory modification, as evidenced by the statistical significance of the squared terms. To gauge the effect size, I bootstrap the four models in Figure 7. At both ends of the x-axis, where protectionist pressure is weaker, the modification rates are similar across the three evidence categories. Consistent with theoretical expectations, legitimate types are less likely to modify their regulatory trade barriers. In the middle range of the x-axis, however, the red line (zero pages of evidence) and the blue line (more than ten pages) lie significantly above the green line (one to ten pages). This too aligns with the theoretical logic: protectionist types are more likely to modify when they submit either very little or very extensive evidence, compared to when they submit a moderate amount. The effect size is substantively large. The gap between the blue and green lines averages about six percentage points, which is considerable given that countries modify their regulatory trade barriers only about six percent of the time on average (Figure 5). In Model 4, the predicted probability of modification rises from 0.0510 to 0.117 when a country submits more than ten pages instead of one to ten pages of evidence—an increase of 129 percent.

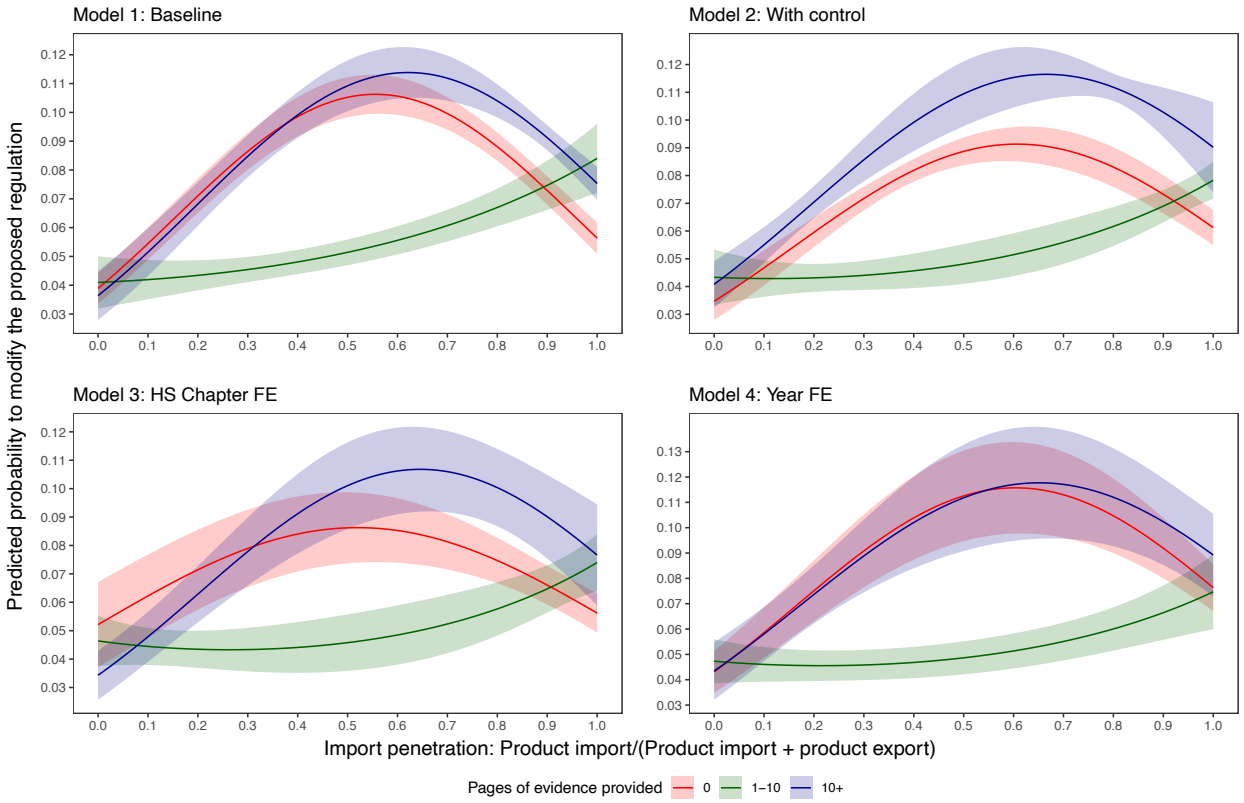


Figure 7: Predicted probability of regulatory modification by international competition and evidentiary volume (Models 1–4)

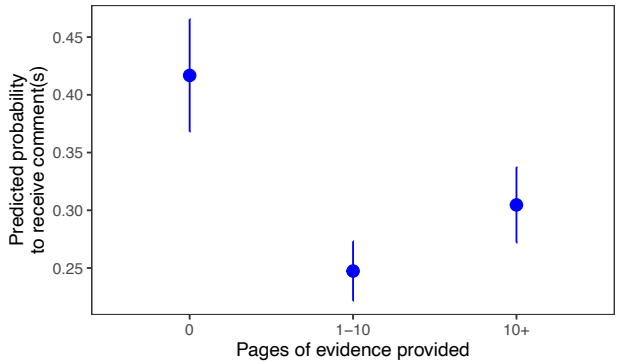


Figure 8: Predicted probability of receiving comments by evidentiary volume.

Consistent with H_3, regulations supported by medium levels of evidence are the least likely to attract comments from other WTO members, as shown in Figure 8. Moving from medium to high evidentiary volume is associated with an increase of approximately 5.7 percentage points—about a 23 percent rise—in the probability that a notification receives a comment, indicating that both low and high levels of evidentiary disclosure invite greater scrutiny than intermediate levels.

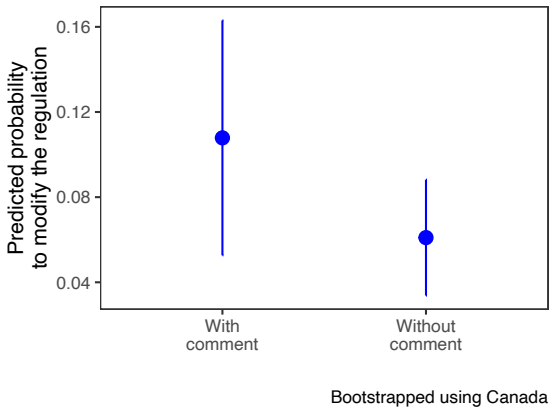


Figure 9: Predicted probability of modification, with and without comments

Consistent with H_4, receiving a comment substantially increases the likelihood that a regulation is modified. As illustrated in Figure 9, for Canada, the predicted probability of modification rises from 6.1 percent in the absence of comments to 10.8 percent when a comment is received—an increase of approximately 4.7 percentage points, or about 77 percent. This pattern indicates that peer feedback within the WTO meaningfully shapes governments' regulatory responses.

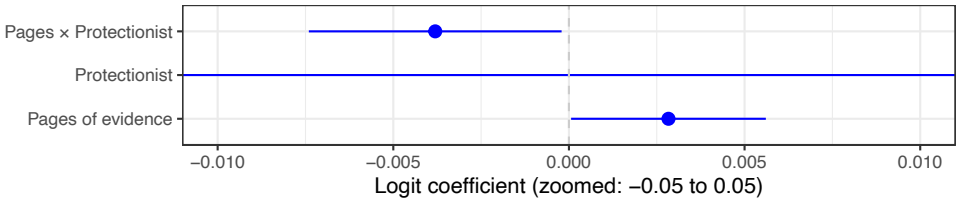


Figure 10: Effect of evidence volume on the probability of receiving a comment, by protectionist status.

Consistent with H_5, increasing evidentiary volume reduces the likelihood that protectionist measures receive comments from other WTO members. In this specification, protectionist status is coded as a binary indicator equal to one when international competition falls between 0.33 and 0.66, capturing product markets characterized by sustained import competition. As shown in Figure 10, a one–standard-deviation increase in pages is associated with an increase of approximately 0.9 percentage points (about 13 percent increase relative to the 5.5 percentage point baseline) in the probability of receiving a comment for non-protectionist measures. In contrast, for protectionist measures, the same increase in evidentiary volume is associated with a decrease of roughly 0.3 percentage points (about 4 percent decrease relative to 5.5 percentage point baseline) in comment probability, indicating that high evidentiary disclosure dampens peer scrutiny when protectionist incentives are strongest.

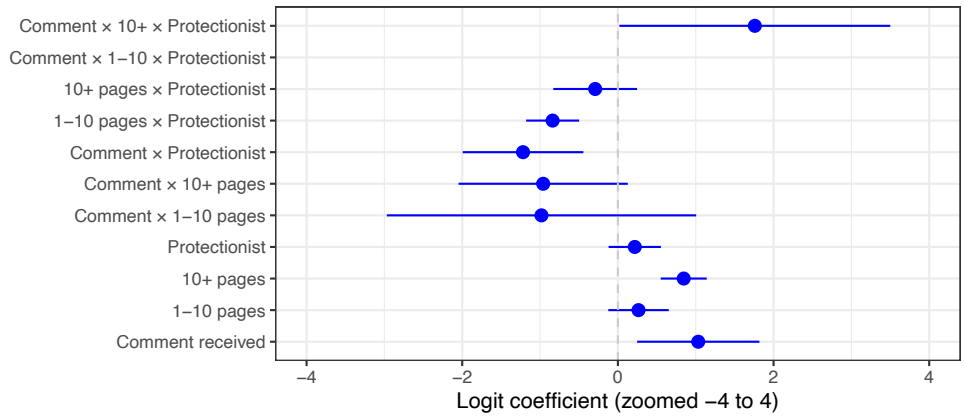


Figure 11: Effect of evidence, protectionist status, comment on modification rate

Consistent with H_6, conditional on receiving a comment, high-evidence protectionist measures are substantially more likely to be modified. As shown in Figure 11, for regulations supported by high evidentiary volume and characterized by strong protectionist pressure, receiving a comment increases the log-odds of modification by 0.614—corresponding to an approximately 84 percent increase in the likelihood of modification. In substantive terms, the predicted probability of modification rises from a baseline of about 6 percentage points to roughly 11 percentage points following a comment, indicating that peer scrutiny is particularly effective in inducing retreat when protectionist measures are backed by extensive documentation.

7 Conclusion

This paper shows that regulatory discipline in the WTO operates primarily through transparency and peer scrutiny rather than formal dispute settlement. By analyzing more than 50,000 TBT and SPS notifications between 2010 and 2024, I demonstrate that governments frequently modify proposed regulatory trade barriers in response to international scrutiny, even in the absence of legal enforcement. The evidence reveals a systematic non-linear pattern. Regulations supported by either very little or very extensive evidence are more likely to be modified, particularly when protectionist incentives are present. In contrast, measures backed by intermediate levels of evidence attract the least scrutiny and are least likely to be commented on. Comments themselves play a decisive role: once a regulation is publicly questioned, the likelihood of modification increases substantially. Moreover, evidentiary disclosure interacts with protectionist incentives in important ways. As evidentiary volume increases, protectionist regulations become less likely to receive comments relative to non-protectionist measures, consistent with strategic over-disclosure aimed at deterring scrutiny. Yet this deterrent effect reverses once scrutiny is triggered: conditional on receiving a comment,

high-evidence protectionist measures are especially likely to be modified.

These findings have important policy implications. Current approaches that rely on unilateral pressure to eliminate non-tariff barriers—such as those pursued by the Trump administration through retaliatory tariffs and aggressive demands for regulatory change—risk overlooking the institutional mechanisms through which regulatory revisions actually occur. Governments are more responsive to how their regulatory justifications are evaluated within multilateral frameworks like the WTO, where documentation functions not merely as procedural compliance but as a reputational signal. As former U.S. ambassador to the EU Anthony Gardner warned, "these things are highly complex and require deep analysis, so if someone like Trump says let's just get rid of them, it's not going to work" (Mackrael, Bade, and Luhnow, 2025). Effective strategies for addressing regulatory trade barriers therefore lie less in coercion than in leveraging transparency, deliberation, and peer scrutiny within multilateral institutions.

References

Bounds, Andy. 2025. "EU plans Donald Trump-style import ban on food." *Financial Times*, February 15. https://www.ft.com/content/e488dc82-e8f6-4a89-8825-0797580211ad.

Busch, Lawrence, Robin Grove-White, Sheila Jasanoff, David Winickoff, and Brian Wynne. 2004. "Amicus Curiae Brief Submitted to the Dispute Settlement Panel of the WTO in the Case of EC-Biotech (WT/DS291, 292 and 293).". Available at: http://www.ecolomics-international.org/biosa_ec_biotech_amicus_academic2_ieppp_lancasteru_coord_0404.pdf (accessed 23 September 2011).

Busch, Marc L. 2000. "Democracy, Consultation, and the Paneling of Disputes under GATT." *Journal of Conflict Resolution* 44(4): 425–446.

Busch, Marc L, and Eric Reinhardt. 2000. "Bargaining in the shadow of the law: early settlement in GATT/WTO disputes." *Fordham Int'l LJ* 24: 158.

Busch, Marc L, and Krzysztof J Pelc. 2015. "Dispute Settlement in the WTO.".

C-SPAN. 2020. "House Ways and Means Committee Hearing on Trade Policy." , June 17. https://www.c-span.org/video/?473040-1/house-ways-means-committee-hearing-trade-policy.

China WTO/TBT National Notification & Enquiry Center. 2022. "Comments from P.R. China on USA Notification G/TBT/N/USA/1528/R1." , July 27. https://www.regulations.gov/comment/EPA-HQ-OAR-2018-0747-0080.

Cho, Sungjoon. 2011. "From control to communication: Science, philosophy, and world trade law." *Cornell Int'l LJ* 44: 249.

de Tannenberg, Valery Laramee. 2016. "The hidden economic benefits of banning plastic bags." *Euractiv*, April 3. https://www.euractiv.com/section/climate-environment/news/the-hidden-economic-benefits-of-the-banning-plastic-bags/.

Downes, Chris. 2012. "The Impact of WTO Transparency Rules: Is the 10,000 th SPS Notification a Cause for Celebration?—A Case Study of EU Practice." *Journal of International Economic Law* 15(2): 503–524.

Hancock, Alice, and Mercedes Ruehl. 2023. "Indonesia and Malaysia freeze trade talks with EU over palm oil." *Financial Times*, May 31.

Herwig, Alexia. 2008. "Whither Science in WTO Dispute Settlement?" *Leiden Journal of International Law* 21(4): 823–846.

Horn, Henrik, Petros C Mavroidis, and Erik N Wijkström. 2013. "In the shadow of the DSU: Addressing specific trade concerns in the WTO SPS and TBT committees." *Journal of World Trade* 47(4).

Howse, Robert. 2000. "Democracy, science, and free trade: risk regulation on trial at the World Trade Organization." *Michigan Law Review* 98(7): 2329–2357.

Howse, Robert, and Joanna Langille. 2023. "Continuity and Change in the World Trade Organization: Pluralism Past, Present, and Future." *American Journal of International Law* 117(1): 1–47.

Jasanoff, Sheila. 2004. "Ordering knowledge, ordering society." In *States of knowledge*. Routledge , 13–45.

Jasanoff, Sheila. 2008. "30 Making Order: Law and Science in Action.".

Karttunen, Marianna B. 2020. *Transparency in the WTO SPS and TBT agreements: the Real Jewel in the Crown*. Cambridge University Press.

Kelley, Judith. 2004. "Who Keeps International Commitments and Why? The International Criminal Court and Bilateral Nonsurrender Agreements." *American Political Science Review* 99(3): 573–589.

Kelley, Judith. 2007. "International actors on the domestic scene: Membership conditionality and socialization by international institutions." *International Organization* 58(3): 425–457.

Lang, Andrew, and Joanne Scott. 2009. "The hidden world of WTO governance." *European Journal of International Law* 20(3): 575–614.

Lange, Jeva. 2023. "The Culture War Comes For Dishwashers." *Heatmap News*, May 10. https://heatmap.news/politics/dishwashers-biden-conservatives-culture-war.

Lester, Simon. 2022. "Where Did "Scientific Basis" Come From in Trade Rules?" *International Economic Law and Policy Blog*, October 2.

Mackrael, Kim, Gavin Bade, and David Luhnow. 2025. "Trump Wants to Rewire Every Other Economy With Trade Onslaught ; The president wants to end so-called non-tariff barriers, including EU regulations that keep out U.S. poultry." *Wall Street Journal*, April 9. https://www.wsj.com/economy/trade/trump-trade-war-economy-goal-c9210091.

Manak, Inveer. 2019. *Enforcing International Trade Law in the World Trade Organization's Committees: Courting Third Party Opinion*. PhD dissertation, Georgetown University.

Meyer, Timothy. 2021. "Cooperating without Sanctions: Epistemic Institutions versus Credible Commitments Regimes in International Law." In *International Law as Behavior*, eds. Harlan Grant Cohen, and Timothy Meyer. Cambridge University Press , 45–73.

Meyer, Timothy. 2022. "The Political Economy of WTO Exceptions." *Washington University Law Review* 99: 1299–1369.

Pauwelyn, Joost, and Weiwei Zhang. 2018. "Busier than ever? A data-driven assessment and forecast of WTO caseload." Oxford University Press.

Pelc, Krzysztof J. 2016. *Making and bending international rules: The design of exceptions and escape clauses in trade law*. Cambridge University Press.

Pevehouse, Jon C. 2005. *Democracy from Above: Regional Organizations and Democratization*. Cambridge: Cambridge University Press.

Rigod, Boris. 2015. *Optimal regulation and the law of international trade: the interface between the right to regulate and WTO law*. Number 18 Cambridge University Press.

Rosendorff, B Peter. 2005. "Stability and rigidity: politics and design of the WTO's dispute settlement procedure." *American Political Science Review* 99(3): 389–400.

Rosendorff, B Peter, and Helen V Milner. 2001. "The optimal design of international trade institutions: Uncertainty and escape." *International Organization* 55(4): 829–857.

Ruehl, Mercedes, Alice Hancock, and Emiko Terazono. 2023. "EU deforestation law triggers ire of its trading partners." *Financial Times*, February 5.

Simmons, Beth A. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Review* 94(4): 819–835.

Sykes, Alan O. 2002. "Domestic regulation, sovereignty, and scientific evidence requirements: a pessimistic view." *Chi. J. Int'l L.* 3: 353.

Tomz, Michael. 2007. *Reputation and International Cooperation: Sovereign Debt across Three Centuries*. Princeton, NJ: Princeton University Press.

Walker, Vern R. 2003. "The myth of science as a neutral arbiter for triggering precautions." *BC Int'l & Comp. L. Rev.* 26: 197.

Wolfe, Robert. 2020. "Reforming WTO conflict management: Why and how to improve the use of 'specific trade concerns'." *Journal of International Economic Law* 23(4): 817–839.

WTO. 2018. "WTO TBT Enquiry Point Guide: Making Transparency Work.".

WTO Appellate Body. 1998. "EC MEASURES CONCERNING MEAT AND MEAT PRODUCTS (HORMONES)." , January 16.

WTO Trade Concerns Database. 2025. "Hong Kong, China; Macao, China; Russian Federation – Restriction on Imports of Aquatic Products from Japan After the Discharge of ALPS Treated Water into the Sea (ID 106).". https://tradeconcerns.wto.org/en/stcs/details?imsId=106&domainId=CMA.

A1 Additional tables

Model:	0. Without interaction (1)
Variables	
Constant	-3.175***
	(0.0594)
1-10 pages	-0.2535***
	(0.0543)
10+ pages	0.1150**
	(0.0471)
Int'l competition	3.042***
	(0.2342)
Int'l competition square	-2.467***
	(0.2199)
Fit statistics	
Observations	35,823
Squared Correlation	0.00736
Pseudo R^2	0.01434
BIC	18,315.9

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 1: Regresssion table for Figure 6

Model:	Baseline (1)	With control (2)	HS chapter FE (3)	Year FE (4)
Variables				
Constant	-3.207***	-7.306***		
	(0.0814)	(0.3046)		
1-10 pages	0.0541	0.2314*	-0.1224	0.0981
	(0.1290)	(0.1342)	(0.2608)	(0.3270)
10+ pages	-0.0691	0.1684	-0.4366	0.0086
	(0.1358)	(0.1389)	(0.3988)	(0.4427)
Int'l competition	3.879***	3.390***	2.079*	3.535***
	(0.3463)	(0.3480)	(1.130)	(0.7699)
Int'l competition square	-3.489***	-2.794***	-2.000*	-2.928***
	(0.3275)	(0.3329)	(1.168)	(0.6735)
1-10 pages × Int'l competition	-3.690***	-3.580***	-2.632*	-3.904***
	(0.5964)	(0.6252)	(1.358)	(1.048)
10+ pages × Int'l competition	0.0634	0.0090	1.679	-0.2259
	(0.5594)	(0.5775)	(1.583)	(1.218)
1-10 pages × Int'l competition squared	4.065***	3.612***	3.049*	3.783***
	(0.5716)	(0.5977)	(1.602)	(0.9829)
10+ pages × Int'l competition squared	0.3159	0.2406	-0.9106	0.3873
	(0.5158)	(0.5351)	(1.845)	(1.039)
Log GDP		0.1489***	0.1648**	0.1627***
		(0.0105)	(0.0663)	(0.0303)
Fixed-effects				
HS Chapter			Yes	
Year				Yes
Fit statistics				
Observations	35,823	33,558	29,905	33,558
Squared Correlation	0.00863	0.01895	0.03552	0.03158
Pseudo R^2	0.01782	0.02801	0.04766	0.05180
BIC	18,293.3	17,585.2	16,175.2	17,282.3

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2: Regression table for Figure 7

Dependent Variables: Model:	Commented (1)	Modified (2)
Variables		
1-10 pages	-0.7766***	
	(0.2818)	
10+ pages	-0.4893*	
	(0.2620)	
Comment received		0.6211**
		(0.2938)
Fixed-effects		
Country	Yes	Yes
Fit statistics		
Observations	29,915	24,986
Pseudo R^2	0.10308	0.18471

Clustered (country) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3: Regression table for Figure 8 and Figure 9

Dependent Variable:	Comment
Model:	(1)
Variables	
No. of pages (scaled)	0.1262**
	(0.0630)
Protectionist	0.0593
	(0.1464)
No. of pages (scaled) × Protectionist	-0.1695**
	(0.0817)
Fixed-effects	
Country	Yes
Fit statistics	
Observations	20,754
Pseudo R^2	0.08727

Clustered (country) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Regression table for Figure 10

Dependent Variable:	Modified
Model:	(1)
Variables	
Comment received	1.033** (0.401)
Medium evidence (1–10 pages)	0.265 (0.198)
High evidence (10+ pages)	0.846*** (0.151)
Protectionist regulation	0.217 (0.171)
Comment × Medium evidence	−0.982 (1.014)
Comment × High evidence	−0.960. (0.555)
Comment × Protectionist	−1.219** (0.396)
Medium evidence × Protectionist	−0.839*** (0.174)
High evidence × Protectionist	−0.292 (0.274)
Comment × Medium evidence × Protectionist	−10.802*** (0.703)
Comment × High evidence × Protectionist	1.760* (0.888)
Fixed-effects	
Country (iso3c)	Yes
Fit statistics	
Observations	17,344
Pseudo R^2	0.178

Clustered (year) standard errors in parentheses
Signif. codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, . $p < 0.15$

Table 5: Regression table for Figure 11

A2 Formal model and hypothesis support

A2.1 Players, timing, and primitives

There are two players. Player 1 (P1) is the regulating government proposing a TBT/SPS measure. Player 2 (P2) is an observing WTO member (or coalition of members) that can engage the measure through the WTO's pre-settlement process.

Types. Nature draws P1's type $\theta \in \{L, P\}$, where L denotes a legitimate regulator and P denotes a protectionist regulator. The type is privately observed by P1. Let $\pi := \Pr(\theta = P)$.

Evidence. P1 chooses evidentiary volume $E \in \mathbb{R}_{\geq 0}$ at the notification stage. Interpreting E empirically, it indexes the volume of attached scientific documentation (pages, studies, annexes). Producing evidence costs $k(E)$.

Timing.

1. Nature draws $\theta \in \{L, P\}$.
2. P1 chooses evidence $E \geq 0$ (notification).
3. P2 observes E and chooses whether to comment $m \in \{0, 1\}$.
4. If $m = 1$, P1 chooses whether to modify $x \in \{0, 1\}$ (modify $x = 1$ vs. resist $x = 0$).
5. If $m = 1$ and $x = 0$, P2 may escalate (e.g., STC/litigation) with probability $\delta \in (0, 1]$.

Beliefs. Upon observing E, P2 forms posterior belief

$$\mu(E) := \Pr(\theta = P \mid E) = \frac{\pi r_P(E)}{\pi r_P(E) + (1 - \pi) r_L(E)},$$

where $r_\theta(E)$ is the equilibrium density over evidence chosen by type θ in the population.

A2.2 Payoffs and behavioral assumptions

P1 receives benefit $v_\theta > 0$ if the measure survives without being modified. Evidence costs $k(E)$ are sunk once produced.

P2's (expected) harm. If the measure remains in place and P1 is protectionist, P2 expects loss $h_P > 0$; if P1 is legitimate, loss $h_L \geq 0$, with $h_P > h_L$. Define $\Delta := h_P - h_L > 0$.

Commenting technology (deterrence). A comment imposes administrative and diplomatic cost to P2 that is increasing in evidentiary volume (processing burden), while the marginal "benefit" of commenting is increasing in expected harm. We summarize this with a reduced-form comment probability

$$q(E) := \Pr(m = 1 \mid E),$$

and assume $q'(E) < 0$ (evidence deters comments). In Section A2.6 we discuss microfoundations.

Escalation and exposure. If $m = 1$ and P1 does not modify ($x = 0$), escalation occurs with probability δ, and if escalated the expected penalty to P1 depends on type and evidence. Let $\rho_\theta(E) \in [0, 1]$ denote the probability of "being found problematic" (or suffering reputational/legal loss) conditional on escalation. Let the penalty magnitude be $S > 0$.

P1 payoffs. If $m = 0$, the measure survives and P1's payoff is

$$u_1(E \mid m = 0, \theta) = v_\theta - k(E).$$

If $m = 1$ and P1 modifies ($x = 1$),

$$u_1(E \mid m = 1, x = 1, \theta) = -k(E).$$

If $m = 1$ and P1 resists ($x = 0$),

$$u_1(E \mid m = 1, x = 0, \theta) = v_\theta - k(E) - \delta\rho_\theta(E)S.$$

P2 payoff. We model P2 as minimizing expected loss. Since hypotheses focus on the incidence of comments and modification, it suffices that P2 prefers to comment when expected harm reduction exceeds comment costs, and to escalate when expected harm exceeds escalation cost.

A2.3 Assumptions (minimal and empirically interpretable)

A1 (Values). $v_P \geq v_L > 0$.
A2 (Evidence cost). $k'(E) > 0$, $k''(E) \geq 0$, and $k(0) = 0$.
A3 (Deterrence). $q'(E) < 0$.
A4 (Exposure in high evidence). $\rho'_P(E) > 0$ and $\rho_P(E) > \rho_L(E)$ for all E in a high-evidence region $\mathcal{E}^H$.
A5 (Return of suspicion). The posterior $\mu(E)$ is high at low evidence, low at medium evidence, and high again at very high evidence; equivalently, $\mu(E)$ is non-monotone with a minimum on a medium-evidence region $\mathcal{E}^M$ and higher values on $\mathcal{E}^L$ and $\mathcal{E}^H$.

Assumption A5 captures the "return of suspicion" logic: low evidence invites doubt about scientific basis; very high evidence can invite doubt about unnecessary burden, inconsistency, or strategic over-documentation. The model is a population game: both types can appear at low, medium, and high evidence levels through heterogeneity in costs or stakes embedded in $r_\theta(\cdot)$.

A2.4 Equilibrium behavior

Modification rule. Given $m = 1$, P1 modifies iff modifying weakly dominates resisting:

$$\begin{aligned} -k(E) &\geq v_\theta - k(E) - \delta\rho_\theta(E)S \\ &\iff \delta\rho_\theta(E)S \geq v_\theta. \end{aligned}$$

Thus define the modification region for type θ:

$$\mathcal{M}_\theta := \{E : \ \delta\rho_\theta(E)S \geq v_\theta\}.$$

Commenting rule (reduced form). Given E, P2 comments with probability $q(E)$. With A3, higher evidence deters comments:

$$\frac{\partial}{\partial E} \Pr(m = 1 \mid E) = q'(E) < 0.$$

P1 expected payoff (ex ante). For type θ, expected payoff from choosing E is

$$\begin{aligned} U_\theta(E) &= (1 - q(E))\,(v_\theta - k(E)) + q(E)\ \max\{-k(E),\, v_\theta - k(E) - \delta\rho_\theta(E)S\} \\ &= v_\theta - k(E) - q(E)\ \min\{v_\theta,\, \delta\rho_\theta(E)S\}. \end{aligned} \tag{1}$$

This expression clarifies the deterrence logic: the expected "penalty" from comment exposure is multiplied by $q(E)$, so high evidence can be attractive when it reduces $q(E)$.

A2.5 Hypotheses as formal implications

We now show that the model supports all six hypotheses as comparative predictions under A1–A5.

H_4 (comments increase modification).

Proposition 1 (H_4). *Suppose $\delta > 0$ and $\rho_\theta(E)S > 0$ for some E. Then for any type θ and evidence E, the probability of modification is weakly higher conditional on receiving a comment:*

$$\Pr(x = 1 \mid m = 1, E, \theta) \;\geq\; \Pr(x = 1 \mid m = 0, E, \theta) = 0.$$

Reason. Modification is only chosen following $m = 1$ in the timing; without a comment there is no incentive (and no move) to modify.

H_{1a} (low evidence increases modification).

Proposition 2 (H_{1a}). *If $\mu(E)$ is higher on the low-evidence region $\mathcal{E}^L$ than on $\mathcal{E}^M$ (A5) and P2's comment probability is weakly increasing in $\mu(E)$ (screening), then measures with low evidence are more likely to be commented on, and therefore more likely to be modified:*

$$E \in \mathcal{E}^L \Rightarrow \Pr(m = 1 \mid E) \textit{ high} \Rightarrow \Pr(x = 1 \mid E) \textit{ high}.$$

Reason. Low evidence increases suspicion about scientific basis, raising comment incidence; by Proposition 1, comments raise modification.

H_{2a} (excessively high evidence increases modification).

Proposition 3 (H_{2a}). *Under A5, if $\mu(E)$ rises again on the high-evidence region $\mathcal{E}^H$ relative to $\mathcal{E}^M$, then high-evidence measures are more likely to be commented on than medium-evidence measures, and thus more likely to be modified:*

$$E \in \mathcal{E}^H \Rightarrow \Pr(m = 1 \mid E) > \Pr(m = 1 \mid E' \in \mathcal{E}^M) \Rightarrow \Pr(x = 1 \mid E) \textit{ higher}.$$

Reason. "Return of suspicion" makes extreme evidence informative again, concentrating comments and subsequent modification.

H_3 (medium evidence attracts the fewest comments).

Proposition 4 (H_3). *If $\mu(E)$ attains its minimum on a medium-evidence region $\mathcal{E}^M$ (A5) and P2 comments when posterior suspicion exceeds a cutoff, then medium-evidence measures attract the fewest comments:*

$$E \in \mathcal{E}^M \Rightarrow \Pr(m = 1 \mid E) \textit{ is minimized}.$$

H_5 (as evidence increases, protectionist measures are less likely to be commented relative to non-protectionist measures).

Proposition 5 (H_5)**.** *Under A1–A3, suppose P-types place greater value on deterrence (A1) and therefore optimally choose higher evidence on average, i.e. the distribution $r_P(E)$ first-order stochastically dominates $r_L(E)$. Then the unconditional incidence of comments is lower for protectionist measures:*

$$\mathbb{E}_{r_P}[q(E)] < \mathbb{E}_{r_L}[q(E)].$$

Moreover, within evidence bins, the marginal effect of evidence on comment probability is negative:

$$\frac{\partial}{\partial E}\Pr(m = 1 \mid E, \theta) = q'(E) < 0,$$

so that increasing evidentiary volume makes comments less likely for protectionist measures as well.

Reason. Since comments are deterred by evidence (A3), types with higher stakes optimally "buy" deterrence by increasing E; hence protectionist measures are less likely to be commented on, especially as evidence rises.

H_6 (conditional on comment, high-evidence protectionist measures modify more).

Proposition 6 (H_6)**.** *Under A4, conditional on being commented, protectionist measures are more likely to fall into the modification region at high evidence:*

$$E \in \mathcal{E}^H \Rightarrow \rho_P(E) \textit{ high} \Rightarrow \delta\rho_P(E)S \geq v_P \Rightarrow x = 1.$$

In particular, if $\rho_P'(E) > 0$ on $\mathcal{E}^H$, then $\Pr(x = 1 \mid m = 1, E, P)$ is increasing in E on $\mathcal{E}^H$.

Reason. High evidentiary volume increases exposure conditional on comment and escalation, making modification the dominant response for protectionist measures in the high-evidence region.

H_{1b} and H_{2b} (protectionist regulation increases modification at low and high evidence).

Proposition 7 (H_{1b} and H_{2b})**.** *Suppose $\rho_P(E) > \rho_L(E)$ (A4) and $\delta\rho_P(E)S \geq v_P$ holds on low- and high-evidence regions (possibly through different mechanisms: low evidence triggers strong peer pressure; high evidence increases exposure). Then protectionist measures are more likely to be modified in both low- and high-evidence regions:*

$$E \in \mathcal{E}^L \cup \mathcal{E}^H \Rightarrow \Pr(x = 1 \mid m = 1, E, P) \geq \Pr(x = 1 \mid m = 1, E, L).$$

Reason. Protectionist measures face larger expected penalties conditional on comment, so they are more likely to retreat (modify) when challenged in the low- and high-evidence regions.

A2.6 Discussion and microfoundations

The reduced-form deterrence assumption $q'(E) < 0$ (A3) can be microfounded in two standard ways without altering the equilibrium logic: (i) commenting requires costly expert attention, and evidence volume raises processing costs; or (ii) high-evidence notifications create a presumption of legal/technical defensibility, reducing the expected marginal impact of a comment. The exposure assumption $\rho_P'(E) > 0$ (A4) captures that, conditional on a comment and escalation, richer dossiers provide more material for scientific and legal scrutiny, increasing the probability that protectionist elements are identified and disciplined.

Together, these ingredients produce the central non-monotonicity: medium evidence is least suspicious and receives the fewest comments (Proposition 4), while both low and excessively high evidence concentrate scrutiny and modification (Propositions 2–3), with protectionist measures strategically buying deterrence through higher evidence (Proposition 5) yet retreating more once challenged at high evidence due to increased exposure (Proposition 6).