

Determining the Appropriate Standard of Review in WTO Disputes

Andrew T. Guzman[†]

Introduction	45
I. The Importance of the Standard of Review	47
II. What We Know About Standard of Review	48
III. Bias vs. Expertise	52
A. Different States, Different Policies	54
B. One World, One Policy	56
IV. Balancing Neutrality and Sovereignty	57
A. De Novo Review (or Close to it)	58
1. <i>The Special Case of Questions of Law</i>	58
2. "All Relevant Factors"	60
3. <i>Objective Evidence</i>	62
B. Intermediate Review: Evidence Involving Judgment and Discretion	64
C. Deferential Review: Science, Risk, and Risk Assessment	68
D. The Special Case of the AD Agreement	73
Conclusion	75

Introduction

The phrase "standard of review," when used in domestic law, normally refers to the aggressiveness with which an appellate court will review a lower court's ruling.¹ A low, or deferential, standard of review implies that a higher court will reverse a lower court's ruling only if the lower court commits a serious error.² A high, or stringent, standard of review implies that a higher court will reverse a lower court's decision if the appellate court finds there to have been any error at all.³

[†] Professor of Law and Director of Graduate Programs, Berkeley Law School. J.D. Harvard Law School, Ph.D. (Economics) Harvard University. Thanks to Joost Pauwelyn and the participants at the *Berger/Cornell International Law Journal* Symposium on Process and Procedure in WTO Dispute Settlement.

1. See Martha S. Davis, *A Basic Guide to Standards of Judicial Review*, 33 S.D. L. REV. 469, 470 (1988) (identifying six distinct standards of review in domestic litigation).

2. See *id.* at 471, 475-81 (identifying six distinct standards of review in domestic litigation and the level of deference each accords).

3. *Id.* The same terminology is also used, in similar fashion, to describe judicial review of legislative or administrative actions. See *Chevron U.S.A., Inc. v. Nat'l Res. Def. Council, Inc.*, 467 U.S. 837, 842-43 (1984) (holding that a court shall apply a deferential standard of review to an agency's reasonable construction of a statute that is silent or ambiguous regarding the issue); Stephen Breyer, *Judicial Review of Questions of Law and Policy*, 38 ADMIN. L. REV. 363, 365-67 (1986) (discussing the principles of law governing judicial standards of review of agency action).

42 CORNELL INT'L L.J. 45 (2009)

The World Trade Organization (WTO), however, defines “standard of review” somewhat differently.⁴ Here, the term refers to the review of national decisions or policies by WTO panels or the Appellate Body (AB).⁵ The distinction between deferential and stringent review remains, but the decision under review is the one made by one of the litigants to the case rather than by a lower court.⁶ This distinction is important to understanding both the existing jurisprudence of the WTO’s Appellate Body (AB) and the normative implications of its approach.

In general, the role of trial and appellate courts is based on their relative strengths and weaknesses. The standard of review also reflects the abilities of each court. In domestic litigation, appellate courts are valued because they bring consistency to lawmaking, because they are thought to be (on average) more qualified jurists, and because they can focus more closely on questions of law.⁷ By contrast, trial courts see much more of the case and are thus better positioned than appellate courts to determine issues of fact.⁸

The WTO should, and to a significant extent does, determine the standard of review applied to national decisions using principles similar to those in domestic litigation. It should also take into consideration the relative capabilities of national authorities and WTO panels. In particular, while WTO panels have the merit of neutrality when reviewing a case, which normally counsels in favor of allowing them to exercise more stringent review, their weakness is similar to that of most reviewing courts: they are less knowledgeable about the facts of the case.⁹ This means that—as compared to a national government—the panel is less familiar with the events that took place, is poorly positioned to evaluate the preferences and attitudes of individuals, and is less skilled at evaluating political, cultural, and other factors relevant to a decision.

4. See Matthias Oesch, *Standards of Review in WTO Dispute Resolution*, 6 J. INT’L ECON. L. 635, 637–39 (2003) (defining standard of review in the context of the WTO).

5. *Id.* at 637. The standard of review can also refer to the review of panel decisions by the Appellate Body, but that is a different issue that is not addressed in this article. See, e.g., Claus-Dieter Ehlermann, *Six Years on the Bench of the “World Trade Court”—Some Personal Experiences as Member of the Appellate Body of the World Trade Organization*, 36 J. WORLD TRADE 605, 619 (2002).

6. See Oesch, *supra* note 4, at 637–38.

7. See, e.g., *Salve Regina Coll. v. Russell*, 499 U.S. 225, 232 (1991) (“Courts of appeals . . . are structurally suited to the collaborative juridical process that promotes decisional accuracy. With the record having been constructed below and settled for purposes of the appeal, appellate judges are able to devote their primary attention to legal issues.”).

8. See, e.g., *Pierce v. Underwood*, 487 U.S. 552, 560 (1988) (stating that due to pretrial activities, a lower court may have insights the record may not convey, and that the only way an appellate court can get full knowledge of the full factual record is to review the entire record).

9. See Appellate Body Report, *European Community—Measures Concerning Meat and Meat Products (Hormones)*, ¶ 115, WT/DS26/AB/R, WT/DS48/AB/R (Jan. 16, 1998) [hereinafter *EC—Hormones AB Report*] (stating that the standard of review “must reflect the balance established in [the SPS] Agreement between the jurisdictional competences conceded by the Members to the WTO and the jurisdictional competences retained by the Members for themselves”).

This article considers how the applicable standard of review should be determined based on the relative strengths of the WTO adjudicatory organs and domestic governments. It explains when review at the WTO should be *de novo*, when there should be an intermediate review, and when the panel or AB should largely defer to the decisions of member states. In the process, it compares these normative claims to existing WTO practice.

I. The Importance of the Standard of Review

WTO agreements impose a broad set of legal obligations on member states with which these states are expected to comply. To judge compliance, WTO panels and the AB must determine how to review states' actions. For example, is a measure supported by sufficient scientific evidence as required by the Agreement on the Application of Sanitary and Phytosanitary Measures (SPS)?¹⁰ Has there been an injury sufficiently serious to justify a safeguard measure? Are two goods "like products"? Is a given measure necessary to protect public morals?

One possible strategy for evaluating state compliance would be to have a WTO panel review all such questions *de novo*, substituting its own judgment for that of the relevant member state. This method would have the merit of putting the key judgment in a neutral party's hands, rather than in the hands of a member state with an interest in the outcome. It would also, however, have a host of disadvantages. For instance, such a rule would move a tremendous amount of decision-making authority to the panel and the AB. These organs would have to determine, for example, the appropriate level of risk that citizens within a member state should have to face before they regulate the importation of a potentially dangerous product. Such an approach would greatly undermine the authority of states to determine their own policies, including in areas long thought to be under domestic control such as health policy, environmental regulation, and labor standards.

Once it is agreed that a *de novo* approach should not be applied in all circumstances, the standard of review question becomes critical. For example, if it is alleged that a member state implemented a safeguard in violation of the Agreement on Safeguards, it is necessary to first determine whether the member state imported a product "under such conditions as to cause or threaten to cause serious injury to the domestic industry that produces like or directly competitive products."¹¹ How should a WTO panel or the AB judge the adequacy of the defendant's investigation?

10. Agreement on the Application of Sanitary and Phytosanitary Measures, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1867 U.N.T.S. 493 [hereinafter SPS Agreement].

11. Agreement on Safeguards art. 2(1), Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, Legal Instruments—Results of the Uruguay Round, 1869 U.N.T.S. 154 [hereinafter Safeguards Agreement].

The defendant state is certain to know more about its own economy and the impact of imports than does a panel or the AB. Moreover, the state has more time and greater resources than WTO tribunals. The state is, therefore, better positioned than WTO judicial organs to determine whether the importation of a product caused injury to the state. Conversely, the defendant state may face domestic political pressures to violate the agreement. This reality raises the possibility that the defendant state will misrepresent its conclusions with respect to injury in order to achieve its own domestic objectives. How one balances the need for objectivity against the informational advantage member states possess has a critical impact on the extent to which WTO obligations constrain states.

II. What We Know About Standard of Review

The Understanding on Rules and Procedures Governing the Settlement of Disputes¹² (DSU) establishes the standard of review in WTO disputes. Article 11 of the DSU states that “a panel should make an objective assessment of the matter before it, including an objective assessment of the facts of the case and the applicability of and conformity with the relevant covered agreements.”¹³ This standard of review applies generally to all WTO disputes unless a more specific rule exists.¹⁴ On its face, the language of Article 11 does not clarify whether a panel or the AB should engage in *de novo* review or the extent to which they should defer to the determinations of national authorities.¹⁵

Specific provisions in the various agreements could have further clarified the appropriate standard of review. With the exception of Article 17.6 of the Anti-Dumping (AD) Agreement, however, the member states declined to include that level of detail.¹⁶ They did not specify a standard of review in the SCM Agreement,¹⁷ the SPS Agreement,¹⁸ the Safeguards Agreement,¹⁹ or any of the other WTO agreements.²⁰

12. Understanding on Rules and Procedures Governing the Settlement of Disputes, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 2, Legal Instruments—Results of the Uruguay Round, 33 I.L.M. 1125 (1994) [hereinafter DSU].

13. *Id.* art. 11.

14. For example, a specific standard of review applies in anti-dumping disputes. See Agreement on Implementation of Article VI of the General Agreement on Tariffs and Trade 1994 art. 17.6, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1868 U.N.T.S. 201 [hereinafter AD Agreement].

15. See Claus-Dieter Ehlermann & Nicolas Lockhart, *Standard of Review in WTO Law*, 7 J. INT'L ECON. L. 491, 495 (2004) (stating that “the standard of ‘objective assessment’ is couched in rather broad terms that do little to provide substantive guidance on the nature and intensity of the review which panels should apply to national measures”).

16. MATTHIAS OESCH, STANDARDS OF REVIEW IN WTO DISPUTE RESOLUTION 7 (2003); see AD Agreement art. 17.6.

17. Agreement on Subsidies and Countervailing Measures, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A, 1867 U.N.T.S. 14 [hereinafter SCM Agreement].

18. See generally SPS Agreement.

19. See generally Safeguards Agreement.

20. OESCH, *supra* note 16, at 6–7.

The question of the appropriate standard of review is closely related to the question of what authority and power the WTO has and what power remains with member states. Member states, for example, retain authority over domestic regulatory policy, but when that policy impacts trade, they may need to turn to Article XX of the GATT²¹ to defend their actions.²² If they do so, a panel or the AB must determine the degree of deference to give to the member when assessing whether a measure satisfies the requirements of Article XX.²³ If, for example, Article XX(d) is at issue, the panel must rule on whether the measure is “necessary” to secure compliance with relevant laws or regulations.²⁴ Granting greater deference to the state’s decisions is equal to increasing its substantive power to impact trade.

As already mentioned, the AD Agreement, in contrast to all other WTO Agreements, includes fairly explicit instructions regarding the standard of review:

- (i) in its assessment of the facts of the matter, the panel shall determine whether the authorities’ establishment of the facts was proper and whether their evaluation of those facts was unbiased and objective. If the establishment of the facts was proper and the evaluation was unbiased and objective, even though the panel might have reached a different conclusion, the evaluation shall not be overturned;
- (ii) the panel shall interpret the relevant provisions of the Agreement in accordance with customary rules of interpretation of public international law. Where the panel finds that a relevant provision of the Agreement admits of more than one permissible interpretation, the panel shall find the authorities’ measure to be in conformity with the Agreement if it rests upon one of those permissible interpretations.²⁵

The standard applied here is clearly deferential to both the factual and legal determinations made by the anti-dumping authorities of states. Under (i), the panel is charged only with determining if the establishment of the facts was proper and if the evaluation was unbiased.²⁶ If so, the panel must accept the conclusions that the state reached.²⁷ Under (ii), the panel must defer to the legal interpretation of the state authorities as long as that interpretation is reasonable.²⁸

21. See General Agreement on Tariffs and Trade art. XX, Oct. 30, 1947, 61 Stat. A-11, 55 U.N.T.S. 194 [hereinafter GATT].

22. Bradley J. Condon, *GATT Article XX and Proximity-of-Interest: Determining the Subject Matter of Paragraphs B and G*, 9 UCLA J. INT’L L. & FOREIGN AFF. 137, 138, 142 (2004).

23. See OESCH, *supra* note 16, at 20 (stating that the issue of the appropriate standard of review to apply arises whenever panels and the AB must interpret and apply domestic law); Steven P. Croley & John H. Jackson, *WTO Dispute Procedures, Standard of Review, and Deference to National Governments*, 90 AM. J. INT’L L. 193, 194 (1996) (observing that GATT dispute settlement procedures have confronted questions of national sovereignty when determining the degree to which an international body should question a national government’s decision).

24. GATT art. XX(d).

25. AD Agreement art. 17.6.

26. *Id.*

27. *Id.*

28. *Id.*

For any other dispute, however, the key language is in Article 11 of the DSU.²⁹ One interpretation of that provision, however, suggests that even Article 11 does not instruct panels or the AB on the standard of review at all. In the *Chile–Agricultural Products* case,³⁰ the AB stated that a panel fails in its obligations under Article 11 if it makes a finding on a matter not before it “because it will thereby fail to accord to a party a fair right of response.”³¹ This denial of a fair right of response is a failure “to accord to [the member] the due process rights to which it is entitled under the DSU.”³² The suggestion is that DSU Article 11 is concerned with providing litigating states certain due process rights rather than with establishing the appropriate standard of review.³³ Because any standard of review—including *de novo* review, complete deference, and anything in between—is consistent with a due process requirement, this interpretation would imply that even the modest guidance that might otherwise be found in DSU Article 11 is not relevant.

The reasoning from *Chile–Agricultural Products* is broadly consistent with the conclusions of the AB in *United States–Hot Rolled Steel* in which the AB addressed the relationship between DSU Article 11 and AD Agreement Article 17.6.³⁴ Notice that if DSU Article 11 is taken to prescribe a particular level of review—if, for example, the words “an objective assessment of the matter before it”³⁵ are taken to require a *de novo* review of facts or law—then there is an apparent tension between DSU Article 11 and AD Agreement Article 17.6.³⁶ At least with respect to an AD dispute,³⁷ one would have to determine which of the two standards applied. In the *Hot Rolled Steel* case, however, the AB stated that there was “no ‘conflict’ between Article 17.6(i) of the *Anti-Dumping Agreement* and Article 11 of the DSU.”³⁸ In the judgment of the AB, then, the deferential standard established by AD Article 17.6 is consistent with the requirement of DSU Article 11.³⁹ This decision implies that the requirement of objectivity in DSU Article 11 is not to be read as demanding a certain standard of review. Instead,

29. See DSU art. 11; OESCH, *supra* note 16, at 98.

30. Appellate Body Report, *Chile–Price Band System and Safeguard Measures Relating to Certain Agricultural Products*, WT/DS207/AB/R (Sept. 23, 2002).

31. *Id.* ¶ 176.

32. *Id.*

33. *Id.*; OESCH, *supra* note 16, at 83.

34. Appellate Body Report, *United States–Anti-Dumping Measures on Certain Hot-Rolled Steel Products from Japan*, ¶ 55, WT/DS184/AB/R (July 24, 2001) [hereinafter *Hot-Rolled Steel AB Report*].

35. DSU art. 11.

36. If these two provisions were in tension, the AD Agreement rule would govern. See *id.* art. 1.2, app. 2 (indicating that to the extent there is a difference between the rules and procedures of the DSU and the rules and procedures set forth in the covered agreements identified in Appendix 2, the rules specified in the covered agreements will prevail).

37. If bilateral attempts between two members to resolve a dispute under the AD Agreement fail, a member may refer the dispute to the Dispute Settlement Body provided for in the AD Agreement. See AD Agreement art. 17.4.

38. *Hot-Rolled Steel AB Report*, *supra* note 34, ¶ 55.

39. See *id.*

it merely requires that there be no bias applied to the adopted standard.⁴⁰

Not all AB decisions, however, share this perspective on DSU Article 11. The AB took a different view in *EC–Hormones*,⁴¹ concluding that DSU Article 11 does indeed prescribe the standard of review that panels are to adopt. The AB declared that “[s]o far as fact-finding by panels is concerned, their activities are always constrained by the mandate of Article 11 of the DSU: the applicable standard is neither *de novo* review as such, nor ‘total deference’, but rather the ‘objective assessment of the facts.’”⁴² The suggestion here is that DSU Article 11 does indeed prescribe a particular standard of review and that both a highly deferential standard and a *de novo* standard would be in conflict with this provision. Even if the interpretation in *EC–Hormones* is correct, that case does not extend beyond ruling out the most extreme forms of review. In this sense the differences between *EC–Hormones*, *Hot-Rolled Steel*, and *Chile–Agricultural Products*, though conceptually significant, have little impact in practice. All three decisions leave open the question of the appropriate standard of review, and all three decisions are consistent with an interpretation that DSU Article 11 provides little or no guidance on the question.

One claim about Article 11 that is likely to be widely accepted is that a “one size fits all” take on the standard of review is incomplete.⁴³ Article 11 requires that panels and the AB make an “objective assessment.”⁴⁴ Objectivity is not only desirable but is also, as I argue later, the defining feature that makes WTO dispute settlement valuable. Objectivity, however, is not sufficient to resolve disputes. The WTO must establish the degree of deference that members’ decisions should receive. In contrast to the objectivity requirement spelled out in DSU Article 11, there is no single “correct” rule as to degree of deference. In some areas, states should be entitled to greater deference, while in other areas more stringent review is necessary if WTO obligations are to be respected.

Existing WTO jurisprudence has implicitly acknowledged that the standard of review must vary depending on the agreement at issue and the matter being decided.⁴⁵ Indeed, even within a single agreement different

40. See *id.* (stating that “it is inconceivable that Article 17.6(i) should require anything other than that panels make an *objective* ‘assessment . . .’”).

41. See *EC–Hormones* AB Report, *supra* note 9.

42. *Id.* ¶ 117.

43. See Ehlermann & Lockhart, *supra* note 15, at 496 (stating that the scope and intensity of the panel’s objective assessment is not the same for every issue, nor in every dispute, and suggesting that the “one size fits all” approach is therefore not practical); see also CATHERINE BUTTON, *THE POWER TO PROTECT: TRADE, HEALTH AND UNCERTAINTY IN THE WTO* 186 (2004) (stating that instead of looking for a non-existent ready-made standard of review, the AB should “balance between regulatory autonomy and international supervision . . . in order to support a revised standard of review”).

44. DSU art. 11.

45. See, e.g., *EC–Hormones* AB Report, *supra* note 9, ¶ 115 (“The standard of review appropriately applicable in proceedings under the SPS Agreement, of course, must reflect the balance established in that Agreement between the jurisdictional competences conceded by the Members to the WTO and the jurisdictional competences retained by the Members for themselves.”).

standards of review are applied depending on the issue in question. This is clear, for example, if one considers some of the obligations within the SPS Agreement. On the one hand, a *de novo* review of the health threat posed in an SPS case would represent a dramatic and stunning intrusion into core areas of domestic authority.⁴⁶ Not surprisingly, the AB has avoided that approach. On the other hand, a stringent review of the requirement that states select the least trade-restrictive measure available is an important check on protectionist impulses and falls squarely within the Dispute Settlement Body's (DSB) mandate to resolve disputes about trade matters.⁴⁷

The lesson from all of this is that both common sense and existing doctrine suggest that DSU Article 11 does not, cannot, and should not prescribe a single standard of review for all cases. It follows that the standard of review must vary depending on the context, and the AB must identify the circumstances in which panels should adopt more or less stringent standards. This variation in the applicable standard of review includes, but is not limited to, distinctions between the standard of review applicable to factual determinations as compared to legal ones.⁴⁸ For questions that are purely matters of WTO law, panels engage in *de novo* review of the judgments made by national authorities.⁴⁹ This approach is required by the DSB's exclusive role in interpreting the WTO agreements.⁵⁰ Granting deference to national authorities would either grant individual states the power to interpret the WTO agreements in a way that would constrain other states (if that interpretation was accepted for future cases) or would create a system in which different states would be able to interpret obligations to suit themselves, creating obligations that vary from state to state. Alternatively, for questions of fact, a more deferential review is appropriate, as discussed later in this article.

III. Bias vs. Expertise

At least three things are clear from the prior discussion about the standard of review as applied by the AB to date. First, the AB has, in practice, significant flexibility in determining the appropriate standard of review. Second, the AB has no choice but to determine the standard of review that it will apply and the standard that the panels must apply. Third, the standard of review applied can vary from issue to issue.

46. BUTTON, *supra* note 43, at 165 (observing that it makes a great "difference to the scope of Members' autonomy whether the panel assesses the existence of a proper scientific justification *de novo*, or whether it exercises . . . deference").

47. See Andrew T. Guzman, *Food Fears: Health and Safety at the WTO*, 45 VA. J. INT'L L. 1 (2004).

48. See *EC—Hormones* AB Report, *supra* note 9, ¶¶ 112–118 (distinguishing the standard of review applicable to factual as opposed to legal disputes).

49. Ehlermann & Lockhart, *supra* note 15, at 498.

50. The DSB has exclusive authority to interpret agreements outside of a formal interpretation by the Ministerial Conference or the General Council. See Marrakesh Agreement Establishing the World Trade Organization art. IX(2), Apr. 15, 1994, 1867 U.N.T.S. 154 [hereinafter WTO Agreement].

The next question is how the AB should determine the appropriate standard of review. Without guidance from the relevant legal texts, the AB has little choice but to do what judicial organs throughout the world do in such situations: base its decisions on relevant policy considerations. To do so, of course, requires that those policy concerns be identified and evaluated.

To start such an analysis it is helpful to ask whether the AB (or a panel) is in a better position than a state to reach a determination on an issue. Or, put differently, will better outcomes be achieved if member states can interpret a given legal rule unilaterally, or is it better to have a panel or the AB second-guess a state's interpretation? Answering this question definitively requires a measure of performance—some way to evaluate outcomes as better or worse. Within the context of evaluating the standard of review, however, no such measure is evident, nor would constructing such a measure be easy. Because the WTO agreements are the product of negotiations among states,⁵¹ it is difficult to speak of an "objective" of those agreements beyond what is included in their texts. It is incorrect to say that there is a goal, for example, of more liberalized trade. The drafters certainly could have provided language that reduced the barriers to trade far beyond the current standards, but they did not do so. Furthermore, the WTO agreements do not charge the WTO judicial organs with the task of achieving objectives that are independent of their language.⁵²

Though the evaluative problems discussed earlier pose difficulties in determining the preferred approach in each instance, it is possible to make relative assessments. That is, it is possible to make comparisons across different types of issues or questions and provide a ranking—ordering the issues from those that should receive the greatest deference to those that should receive the least. We can identify factors that should push toward more or less deferential review. In doing so, we can then describe the variables that should influence the decision and provide guidance on the appropriate standard a judicial body should apply.

There is little disagreement that interesting standard of review issues involve a balancing of the authority of panels and the AB against the sovereign rights of states to set their own policy.⁵³ No one would disagree, for example, with the claim that sovereign states are entitled to establish their own standards with respect to environmental, labor, and health policies. When the impact of these policies on trade becomes serious enough, how-

51. INFO. & MEDIA RELATIONS DIV., WORLD TRADE ORG., UNDERSTANDING THE WTO 9-10 (3d ed. 2007) (stating that the WTO was created by negotiations and that everything it does is the result of negotiations).

52. See Richard H. Steinberg, *Judicial Lawmaking at the WTO: Discursive, Constitutional, and Political Constraints*, 98 AM. J. INT'L L. 247, 250-51 (2004) (stating that to help ensure the WTO dispute settlement system would not shift the members' rights and responsibilities the DSU says that the AB cannot add or diminish the rights and obligations covered in the agreements).

53. See Ehlermann & Lockhart, *supra* note 15, at 494-96.

ever, WTO obligations may constrain the policy options of the state.⁵⁴

In the first instance, this balancing of interests is, and should be, done through the substantive rules of the WTO. The relevant rules permit measures that do not nullify or impair the benefits that other member states enjoy.⁵⁵ For example, a state is free to tax economic activity as it wishes, but the rules constrain the state's ability to do so in a way that discriminates against foreign products. In the second instance, interpretive questions about the meaning of the words exist in the WTO agreements. It is only after these questions of interpretation are addressed that the standard-of-review issue comes into play. In particular, to what extent should a panel or the AB review the compliance of measures enacted by member states with WTO obligations? Moreover, to what extent should a panel or the AB evaluate local priorities? In what circumstances should a panel or the AB apply its own judgment to the question of how alternative policies will work within a member state?

A. Different States, Different Policies

To speak coherently about the appropriate standard of review, we must first have a sense of the trade-off between leaving greater discretion in the hands of states and having panels and the AB review state activities with greater care. What are the features of panels and the AB that make them more appropriate places in which to locate decisions about state conduct? It seems that the key feature of these judicial organs is their neutrality. Unlike member states, these judicial organs have no reason to favor the goals or ambitions of one state over those of another.⁵⁶

Member states, conversely, are familiar with the concerns and priorities of their own citizens and have a better understanding of domestic markets and institutions. In addition, they are better positioned to carry out all manner of information gathering and investigation than are panelists or AB members. States can devote greater resources to this sort of activity, will normally have more time to gather and evaluate information, and can more easily make use of relevant experts.

The dilemma inherent in selecting a standard of review is familiar. On the one hand, it is important that trading rules be applied to all states in a consistent and evenhanded way. Doing so encourages trade and discourages the adoption of costly protectionist measures. On the other hand, the AB and panels must respect domestic preferences and priorities.

To illustrate the problem, consider the facts from the *EC–Hormones*

54. See Lori M. Wallach, *Accountable Governance in the Era of Globalization: The WTO, NAFTA, and International Harmonization of Standards*, 50 U. KAN. L. REV. 823, 827 (2002).

55. See DSU art. 3.3; see also GATT art. XXIII.

56. See Joost Pauwelyn, *The Transformation of World Trade*, 104 MICH. L. REV. 1, 25–26 (2005) (noting the impartiality of the WTO judicial body). There are a range of concerns one might have about the impartiality of judges or the objectives they seek to pursue. At the WTO, one might also be concerned that panels and the AB are themselves political players within the WTO system and cannot avoid having their decisions affected by the identity of the parties.

case.⁵⁷ At issue in that case were European Community (EC) measures excluding meat treated with certain hormones from its market.⁵⁸ The United States (among others) adopted a different policy position that permitted such meat to be sold in the United States.⁵⁹ One can easily imagine this difference in policies arising even in the absence of trade; that is, these two jurisdictions could reasonably adopt different policies, even if trade issues motivated neither policy.⁶⁰ Furthermore, as long as domestic concerns unrelated to trade motivate these policies, and as long as trade impacts are minimized, there are good reasons to permit these different policies to remain in place.⁶¹

Assuming that trade concerns play no role in this sort of policy decision, it is difficult to provide a compelling argument for why any international tribunal, including one at the WTO, should have second-guessed the policy priorities of Europe in this case. In the absence of a protectionist motive, these different policy preferences reflect differences in risk tolerance or, perhaps, in the interpretation of scientific evidence. Such differences should be accepted and, indeed, encouraged. Individuals may not only have varying preferences regarding, for example, their willingness to accept risk, but they may also be differently situated, causing them to view trade-offs differently.⁶² For example, all else equal, a poor country will likely tolerate larger health risks in exchange for economic gains than will a rich country.⁶³ Furthermore, heterogeneity in rules spurs healthy debate about policies and values and encourages innovation and experimentation, all of which add value to the global community.⁶⁴ Most importantly, because domestic governments are better able to craft policies that suit their citizens' interests than is an international tribunal, the decision about the relevant rules should be left with those governments.⁶⁵

National governments have an additional advantage over the judicial organs of the WTO—they are better at collecting and analyzing data. For example, they are better positioned to determine how best to gather infor-

57. See generally *EC—Hormones* AB Report, *supra* note 9. This example is only intended as an illustration. Other real or hypothetical cases face the same issues.

58. *Id.* ¶ 2.

59. See Phoenix X.F. Cai, *Between Intensive Care and the Crematorium: Using the Standard of Review to Restore Balance to the WTO*, 15 *TUL. J. INT'L & COMP. L.* 465, 507-08 (2007).

60. Guzman, *supra* note 47, at 11, 15-17 (noting that two countries may employ different policies based on domestic values and attitudes, and that the WTO system does not acknowledge legitimate reasons for divergent state preferences).

61. *Id.* at 11-15 (describing the reasons why allowing competing domestic policies is preferable).

62. *Id.* at 12.

63. *Id.* (noting that a poor country may accept a greater risk for less expensive food, while a wealthy country would not).

64. See Paul B. Stephan, *The Futility of Unification and Harmonization in International Commercial Law*, 39 *VA. J. INT'L L.* 743, 792-93 (1999) (arguing that development on national scales allows for mini-laboratories to provide diverse options, while having one international system of development would eliminate innovation).

65. One could add issues of legitimacy to the list of reasons why domestic governments are better situated to make domestic policy. See Guzman, *supra* note 47, at 14.

mation, what evidence can be obtained, how to interpret it, and what methodologies to use.⁶⁶ They are better equipped both because they are more familiar with local conditions and because they have greater capacity and more time to undertake the evaluation of information.⁶⁷

B. One World, One Policy

Despite the advantage that domestic governments have in evaluating local concerns and enacting measures to address those concerns, not all decisions—not even all decisions that implicate traditionally domestic concerns—should be left to domestic authorities alone. States have entered into the WTO system because they value a single set of rules governing trade.⁶⁸ They understand that without such rules, certain incentives can tempt states to adopt policies that needlessly undermine the trading system and that prove costly to all states.⁶⁹ In the DSU, member states made it clear that they believe the WTO system is stronger and more valuable with a system of binding dispute resolution.⁷⁰

Put more bluntly, the dispute settlement system discourages states from adopting measures for protectionist purposes.⁷¹ The leaders of democratic states answer to their own citizens and thus seek to advance state objectives rather than those of the international trading system.⁷² Therefore, although states have collectively agreed to abide by the WTO agreements, panels and the AB help to prevent them from ignoring their trade commitments.

The neutrality of the dispute settlement system is the key feature that allows it to opine on states' behavior. Because the system does not have a bias in favor of any litigant, it has credibility when resolving disputes.⁷³ Its neutrality also implies that its interpretation and application of the language contained in the WTO agreements is more reliable than interpretations by litigants acting out of self-interest.⁷⁴

This neutrality is the closest we can come to an objective metric of performance for alternative standards of review. Whatever standard is applied, it is important that it promote actions that conform closely to the

66. See *id.* at 23 (stating that WTO dispute regulation organs “lack the resources to carry out their own investigation in anything more than a cursory way”).

67. See *id.* at 21–24.

68. See generally WTO Agreement.

69. See, e.g., GATT art. XXIII.

70. DSU art. 3.2.

71. Alan O. Sykes, *Domestic Regulation, Sovereignty, and Scientific Evidence Requirements: A Pessimistic View*, 3 CHL. J. INT'L L. 353, 353 (2002) (stating the reciprocal commitments resulting from the WTO negotiations reduce or eliminate traditional methods of protectionism).

72. Guzman, *supra* note 47, at 15.

73. See, e.g., DSU art. 17.3 (providing that the AB will be comprised of persons unaffiliated with any government who do not have a conflict of interest and that the membership should be broadly representative of the membership of the WTO); see also Guzman, *supra* note 47, at 15–16.

74. See, e.g., Guzman, *supra* note 47, at 14–17 (giving an example of how the WTO's neutrality can lead to a better application and interpretation of the WTO agreements).

requirements of the WTO agreements. The obvious problem is that we often cannot discern what the text of the agreements requires without having a panel and the AB interpret the agreements' text. Nevertheless, a tribunal's unbiased ruling is, all else equal, preferred to a biased one.

Notice that the emphasis on bias is slightly different from the emphasis on uniformity and consistency in application of WTO law that exists in some writing on standard of review.⁷⁵ Though there is reason to want uniformity and consistency in some instances, these desiderata yield neither a policy-based rationale for a particular standard of review⁷⁶ nor an account of existing practices.⁷⁷

Returning to the *EC–Hormones*⁷⁸ example, if we assume that Europe had protectionist reasons for excluding hormone-treated beef, then the WTO was justified in reviewing this protectionist policy.⁷⁹ If the EC can benefit its farmers by excluding hormone-treated beef, for instance, it will be biased toward that policy and may be tempted to find a non-protectionist justification for the policy, such as an alleged concern for the health consequences of the beef.

IV. Balancing Neutrality and Sovereignty

Selecting an appropriate standard of review involves a balancing of the advantages of neutrality against those of state sovereignty. There is no “one size fits all” way to carry out this balance.⁸⁰ Even within an individual WTO agreement the balance will depend on the particular issues at stake.⁸¹ Indeed the interpretation of the substantive law at issue may influence the appropriate standard of review applied in the context of a particular rule.

Consider, for example, the meaning of the term “like products” in GATT Article III:2.⁸² Existing AB jurisprudence provides that a determination of whether two products are “like” involves a consideration of many factors, including the end uses of the product, product characteristics, consumer tastes and habits, the nature and quality of the product, and the

75. See, e.g., OESCH, *supra* note 16, at 30–31.

76. For example, what areas should have uniformity? Why those areas?

77. For example, why require uniformity in the facts necessary to claim injury, but not require uniformity in weighing the merits of scientific evidence?

78. *EC–Hormones* AB Report, *supra* note 9.

79. See Robert Howse, *Democracy, Science, and Free Trade: Risk Regulation on Trial at the World Trade Organization*, 98 MICH. L. REV. 2329, 2329–40 (2000) (stating that SPS provisions apply even to nondiscriminatory regulations that would not violate GATT provisions and thereby implying that even if there is a possibility that a state's trade regulation may be discriminatory, the WTO would review it as it did in the beef hormones dispute).

80. See Ehlermann & Lockhart, *supra* note 15, at 496 (suggesting that a “one size fits all” approach would not work well).

81. See Mads Andenas & Stefan Zleptnig, *Proportionality: WTO Law: In Comparative Perspective*, 42 TEX. INT'L L.J. 371, 395–96 (2007).

82. GATT art. III:2. The term “like products” appears in various other places in the WTO agreements. See, e.g., AD Agreement art. 3.1. For present purposes, however, it is sufficient to focus on this one occurrence.

tariff classification.⁸³ Applying this test inevitably requires detailed knowledge about the domestic market of the relevant member state—knowledge that may be difficult for a panel to quantify or verify.⁸⁴ Therefore, a state may have an advantage over a panel in evaluating whether two products are “like,” which warrants at least some deference to the state’s judgment regarding the relevant domestic factors.⁸⁵ If, however, the AB chooses to rely on a straightforward economic measure—such as the elasticity of substitution between the two products—to determine whether two products are “like,” a panel could easily verify the information needed to reach a conclusion. Because the elasticity of substitution can be calculated using familiar techniques and standard economic data, the information is as readily available to a panel as it is to the member state. In this circumstance, the member state has little informational advantage over the panel, while the panel still has the advantage of being unbiased. Thus, the case for deference would be considerably weaker if the AB were to use this economic measure rule instead of the existing multi-factor test.

This article now considers the normative case for alternative levels of review. Because discussion of the applicable standard of review is necessarily somewhat case specific, what follows is an attempt to provide some useful guidelines for determining the standard of review. In a specific case the general conclusions presented here can be combined with more detailed analysis of the particulars of that case.

A. De Novo Review (or Close to it)

1. *The Special Case of Questions of Law*

The case for de novo review, or something close to it, is at its strongest where sovereignty concerns are modest but neutrality is important.⁸⁶ The standard of review applicable to purely legal matters is the most obvious example.

The WTO itself is the key source of expertise with respect to pure questions of law.⁸⁷ The task of deciding a question of law is really one of interpreting the meaning of the text found in a covered agreement.⁸⁸ Though the legal question must arise within some factual context, the judicial body need not understand any specific information about the enacting country, nor does it need to know of issues unique to that country’s priori-

83. Appellate Body Report, *Japan—Taxes on Alcoholic Beverages*, at 20–21, WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R (Oct. 4, 1996) [hereinafter *Japan—Alcoholic Beverages* AB Report].

84. See *id.* (stating that given the case-by-case analysis that interpreting “like products” requires, panels can only apply their best judgment, and this will involve an element of discretionary judgment).

85. Oesch, *supra* note 4, at 640.

86. See Guzman, *supra* note 47, at 17.

87. See Ehlermann & Lockhart, *supra* note 15, at 496 (stating that one of a panel’s basic activities is to make findings that are purely legal in nature).

88. See *id.* at 497 (observing that the AB must follow the customary rules of interpretation of public international law, as codified in the Vienna Convention, which require the treaty interpreter to seek out the ordinary meaning of the text).

ties or concerns or how its markets function.⁸⁹

There are, however, important systemic issues relating to how the WTO system will function, how the various obligations within the WTO system affect one another, and the content of those obligations for every member state. The characteristics of national authorities are not well suited for addressing these pure questions of law.⁹⁰ Panels and the AB, however, are unbiased and have the appropriate expertise to resolve such questions.⁹¹ As a normative matter, then, panels and the AB should decide such questions of law on a *de novo* basis.⁹² The practice of the WTO's judicial organs is consistent with this approach: purely legal determinations receive *de novo* review.⁹³

Panels and the AB can strike the appropriate balance between the sovereignty of member states and the benefits from uniformity and neutrality within a *de novo* standard of review. This is so because the substantive rules, rather than the standard of review, primarily determine this crucial balance.⁹⁴ In making its interpretation, a panel or the AB can take the importance of sovereignty into account, which can influence the interpretation of the relevant legal rule. Without *de novo* review, or a similar type of review, uniformity would be impossible and WTO obligations would mean different things in different countries.⁹⁵ Not only would this create a confusing legal framework, but it would also upset the balance of concessions that member states made when negotiating the WTO agreements.⁹⁶

89. See *id.* (indicating that panels and the AB must reach their own conclusions about legal conclusions regardless of the nation's application of law).

90. *Id.* at 498 (arguing that allowing individual WTO members to interpret WTO law would lead to loss of uniformity that would undermine the rule-based system because the obligations that members assume would differ from member to member).

91. See Oesch, *supra* note 4, at 643 (arguing that the "expertise argument" favors panels and the AB regarding questions of law because these judicial organs have expertise and institutional knowledge in WTO law and practice).

92. See OESCH, *supra* note 16, at 174 ("[B]oth panels and the Appellate Body have constantly engaged in a *de novo* standard of review of WTO law.").

93. See Ehlermann & Lockhart, *supra* note 15, at 519-20 ("If the panel is reviewing a purely legal issue, the panel conducts an original or *de novo* interpretation . . ."); Holger Spamann, *Standard of Review for World Trade Organization Panels in Trade Remedy Cases: A Critical Analysis*, 38 J. WORLD TRADE 509, 553 (2004) ("[T]he specialists for WTO law are, after all, the Panels and especially the AB . . .").

94. See Ehlermann & Lockhart, *supra* note 15, at 494-97 (noting that panels and the AB must balance the power between WTO members and the WTO by making an "objective assessment" of the matter before them, which—in the case of legal interpretations of WTO agreements—requires them to apply customary rules of interpretation of public international law).

95. *Id.* at 498.

96. *Id.* (stating that the dispute settlement system seeks to preserve the WTO members' balance of rights and obligations by providing for uniform interpretation of WTO law, and that allowing members discretion to interpret WTO law would undermine the core objectives of the rule-based system). It is worth noting at this point that the AB has repeatedly stated that the standard of review applied should not be *de novo*. See, e.g., Appellate Body Report, *United States—Transitional Safeguard Measure on Combed Cotton Yarn from Pakistan*, ¶ 74, WT/DS192/AB/R (Oct. 8, 2001) [hereinafter *US—Cotton Yarn AB Report*] ("[P]anels must examine whether the competent authority has evaluated all relevant factors; they must assess whether the competent authority has examined all the

These arguments in favor of *de novo* review of purely legal questions are consistent with the AB's practice. In the *EC–Hormones* case, the AB rejected the European Community's call for a deferential standard of review with respect to legal rules.⁹⁷ The AB explained that Article 3.2 of the DSU instructs panels to apply customary rules of interpretation of public international law.⁹⁸ The accepted sources for these customary rules are Articles 31 and 32 of the Vienna Convention on the Law of Treaties.⁹⁹ Under these provisions, the treaty interpreter is obliged to interpret the treaty "in good faith in accordance with the ordinary meaning" of the text.¹⁰⁰ Elaboration of this obligation in the Vienna Convention leaves no room for interpretation based on the views of signatories as expressed after the treaty is entered into force. As such, the rules of interpretation of customary law do not afford panels or the AB discretion to defer to member states' views.

2. "All Relevant Factors"

Panels and the AB are frequently called upon to review an investigation by a domestic government.¹⁰¹ An investigation is required, for example, prior to the implementation of safeguards measures, anti-dumping measures, countervailing duties, and SPS measures. When such a case finds its way to a panel, the review of the measure includes two separate elements.¹⁰² The first is the form of the review by domestic authorities.¹⁰³ Did they take into account the right evidence? Did they provide appropriate procedures? And so on. The second stage of the review concerns the substantive question of whether the measure is justified on the merits.¹⁰⁴

pertinent facts and assessed whether an adequate explanation has been provided as to how those facts support the determination; and they must also consider whether the competent authority's explanation addresses fully the nature and complexities of the data and responds to other plausible interpretations of the data. However, panels must not conduct a *de novo* review of the evidence nor substitute their judgement [sic] for that of the competent authority."). The reluctance to conduct a *de novo* review is understandable, but as this reluctance is stated in the cases it can be misleading. Though the AB does not explicitly frame the standard of review in these terms, panels are to engage in a review of purely legal questions and purely objective facts, as already discussed, in a *de novo* fashion. See Ehlermann & Lockhart, *supra* note 15, at 496–98. In this context, the AB's statements are better understood to mean that matters of judgment should not be reviewed *de novo*. This interpretation is consistent with the normative claims made in this article.

97. *EC–Hormones* AB Report, *supra* note 9, ¶¶ 118–119.

98. *Id.* ¶ 118.

99. Ehlermann & Lockhart, *supra* note 15, at 497; see Vienna Convention on the Law of Treaties arts. 31–32, *opened for signature* May 23, 1969, 1155 U.N.T.S. 331, (entered into force Jan. 27, 1980) [hereinafter Vienna Convention].

100. Vienna Convention art. 31(1).

101. See, e.g., *EC–Hormones* AB Report, *supra* note 9, ¶ 2 (calling upon the AB to review the EC's assessment of the risk of importing meat treated with certain hormones to determine whether the EC had complied with the requirement under the SPS Agreement).

102. See Alexis Goh, *The WTO Dispute Settlement System*, 2001 AUSTL. INT'L L.J. 208, 230 (2001).

103. *Id.*

104. *Id.*

Though the required investigation differs somewhat across the various WTO agreements, the first stage tends to include more formal and procedural requirements and the second tends to focus more on substantive questions.¹⁰⁵ The AB makes this distinction in *US–Lamb*, a case concerning safeguards measures.¹⁰⁶ In that case the AB referred to the first element as the “formal” review and considered whether domestic authorities had examined “all relevant factors” as required under Article 4.2(a) of the Safeguards Agreement.¹⁰⁷ The AB referred to the second element as the “substantive” review, which deals with the defendant’s reasons for its safeguard determination.¹⁰⁸

Although the *US–Lamb* case is a safeguards case, the distinction between formal and substantive review is more widely applicable. The explicit obligation to review all relevant facts, for example, is present in the Safeguards and AD Agreements and has been imported into other areas.¹⁰⁹ Even beyond the requirement to review all relevant facts, domestic governments are frequently required to give notice to other members, to provide opportunities for other states to present evidence, and to explain their decisions.¹¹⁰ These, and a host of other requirements present in various covered agreements, can also be categorized as “formal” requirements.

If one considers the balance between the expertise of member states and the neutrality of panels and the AB, how should the WTO review these formal matters? In practice, panels and the AB have adopted a stringent standard of review on the question of whether all relevant facts have been considered.¹¹¹ Is this consistent with the balancing of members’ expertise at the domestic level and the WTO’s neutrality?

The distinction between formal and substantive review found in *US–Lamb*, including its extension to other formal questions, is useful in addressing this question because it identifies issues that raise important concerns about interpretational bias but that can be resolved without significant domestic expertise.¹¹² Consider again the “all relevant factors”

105. This distinction is imperfect, of course. In the SPS Agreement, for example, it is often difficult to separate the “formal” requirement that a measure be based on a risk assessment from the substantive question of how the applicable science is evaluated. See SPS Agreement art. 3.3. Despite these problems, the distinction in the text is useful to frame the standard of review inquiry.

106. Appellate Body Report, *United States–Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia*, ¶ 103, WT/DS177/AB/R, WT/DS178/AB/R (May 1, 2001) [hereinafter *US–Lamb AB Report*].

107. *Id.* ¶¶ 103–104.

108. *Id.* ¶ 103. The first WTO case to discuss the standard of review to be applied was *US–Underwear*, in which the panel concluded that there must be a review of “all relevant facts” and an adequate explanation of how the “facts as a whole supported the determination made.” Panel Report, *United States–Restrictions on Imports of Cotton and Man-Made Fibre Underwear*, ¶ 7.13, WT/DS24/R (Nov. 8, 1996) [hereinafter *US–Underwear Panel Report*].

109. See, e.g., *US–Underwear Panel Report*, *supra* note 108, ¶¶ 1.1, 7.13. (addressing the Agreement on Textiles and Clothing).

110. See, e.g., SCM Agreement.

111. OESCH, *supra* note 16, at 119.

112. *US–Lamb AB Report*, *supra* note 106, ¶ 104.

requirements.¹¹³ Although it is perhaps imaginable that the relevance of a particular fact can turn on details unique to the domestic context, this is likely to be the exception rather than the rule. More commonly, relevance will be clear to an observer, even if that observer is a WTO panel.¹¹⁴ In any event, if a particular fact is not relevant, the member state can nonetheless err on the side of caution and consider it. The other formal requirements—notice requirements, publication requirements, and so on—are even less likely to involve domestic expertise. Thus in these areas, domestic governments have at most a very modest informational advantage over a panel or the AB.

These same issues, however, raise critical questions of bias. By omitting certain relevant facts, a member state can make a decision appear defensible and appropriate when, in fact, it is not.¹¹⁵ Thus, a failure to consider all facts both allows a member state to reach its preferred conclusion—even when that conclusion is not justified once all relevant facts are taken into account—and makes it more difficult for a panel, the AB, or other states to understand the context in which a member state adopted a trade measure. Similarly, a failure to publish information promptly, to allow the submission of evidence, and to explain a state's actions can place imports at a disadvantage, consequently achieving a protectionist objective.¹¹⁶

In light of the above, the adoption of a stringent standard when dealing with formal review seems quite sensible. There is little loss of expertise: the question of relevance is normally not something that differs from country to country, and there is a significant concern of bias if states have the discretion to make determinations about the relevance of particular facts.

3. *Objective Evidence*

Beyond the question of whether all relevant facts have been considered, there can be disputes about the facts themselves—i.e., which alleged facts are true and which are not. In many circumstances, the member state will have a significant advantage in identifying relevant facts. In the case of “raw facts,” meaning facts that can be verified by the panel, there is a strong case for *de novo* review. Indeed, the case of raw facts is an example

113. See, e.g., *id.*

114. See Safeguards Agreement art. 4.2(a) (providing a list of factors that the AB or panel should consider, including “the rate and amount of the increase in imports of the product concerned[,] . . . the share of the domestic market taken by increased imports, changes in the level of sales, production, productivity, capacity utilization, profits and losses, and employment”). Given that the Safeguards Agreement provides a list of factors, the relevance of these factors will be clear within each context.

115. See, e.g., *US—Cotton Yarn AB Report*, *supra* note 96, ¶ 7.35.

116. See Safeguards Agreement art. 3.1 (requiring that an investigation by competent authorities with respect to a safeguard measure include “reasonable public notice to all interested parties . . . or other appropriate means in which . . . interested parties could present evidence and their views” and requiring authorities to “publish a report setting forth their findings and reasoned conclusions[,]” which suggests that not following these requirements would result in protectionism).

of a more general category of evidence that panels and the AB must review. If an issue is to be determined based on objective and verifiable criteria or evidence, there is a stronger case to be made for a fulsome review by a panel or the AB.

For much of the evidence used in disputes, this point is sufficiently clear that in practice the issue does not even arise. For example, there is no dispute that a panel should examine the pattern of imports in a safeguards case to determine if those imports have increased or decreased.¹¹⁷ If the available data shows that imports have risen, a panel will not accept claims from the defendant that they have fallen. Moreover, the question of how to determine whether imports have fallen is one for which panels or the AB show little deference.¹¹⁸

The most obvious category of objective evidence is economic data. There can, of course, be questions of how the data was collected and whether or not it is reliable; but once a source of data is agreed upon, panels and the AB should, and do, engage in a high level of review.¹¹⁹

Notice that here we are discussing only the evaluation of existing and accepted data. Questions of how data should be collected or how reliable it is are less objective, and domestic governments may have expertise in these areas.¹²⁰ For example, local conditions such as the frequency of false reporting and the portion of the local economy included in economic statistics can affect data collection.¹²¹ There may also be disagreements about alternative sources of data with some observers (or member states) favoring one source and others favoring another. Here, domestic governments are likely to have greater expertise than a panel or the AB in evaluating the available alternatives.

For these reasons, the questions of how data is collected and which data is used are less objective than the simple reading of the data once it is accepted. As such, the former questions should be reviewed less stringently than the latter. This approach is broadly consistent with existing practice at the WTO. In *US–Combed Cotton Yarn*, for example, the panel stated that “[i]t is not for panels to prescribe precise methodologies for information gathering and verification. In our view, the US methods used in this proceeding are not unjustifiable, even though we recognize that

117. See *id.* arts. 2.1, 4.2(a).

118. See, e.g., Appellate Body Report, *Argentina–Safeguard Measures on Imports of Footwear*, ¶¶ 129–131, WT/DS121/AB/R (Dec. 14, 1999) [hereinafter *Argentina–Footwear AB Report*] (stating the stringent requirement that the increase in imports must be “recent enough, sudden enough, sharp enough, and significant enough, both quantitatively and qualitatively, to cause or threaten to cause ‘serious injury’”).

119. OESCH, *supra* note 16, at 120.

120. See AD Agreement art. 17.6(i).

121. See U.N. DEP’T OF ECON. & SOC. AFFAIRS, INTERNATIONAL MERCHANDISE TRADE STATISTICS COMPILERS MANUAL at 43–45, U.N. Doc. ST/ESA/STAT/SER.F/87, U.N. Sales No. E.02.XVII.17 (2004), available at http://unstats.un.org/unsd/publication/SeriesF/seriesf_87e.pdf (providing suggestions for increasing data quality control).

there may be other approaches.”¹²² Notice how deferential the language is here: there is no indication that the United States must demonstrate that its method is the best available. It is sufficient that the method is “not unjustifiable” or, as stated in the same paragraph of the ruling, the party disputing the data must show that it is “inherently untrustworthy.”¹²³

The prior subsection addressed the question of how to review whether the domestic authorities have examined all relevant factors in making a determination. The argument for a stringent review in such cases relies on the fact that the question of relevance is fairly objective, meaning that there is little local expertise involved.¹²⁴

For similar reasons, it makes sense for panels or the AB to also apply a stringent standard of review to most issues of transparency and most procedural requirements. There is little to be gained from transnational diversity with respect to these issues, and it is relatively easy for a panel or the AB to determine what is or is not being done by a member state.¹²⁵

Other categories of questions surely exist, but the point should be clear: objective evidence for which there is little domestic expertise and on which there will normally be broad agreement should be subject to a high level of review.

B. Intermediate Review: Evidence Involving Judgment and Discretion

It is often remarked that the AB has taken a more deferential position toward the review of factual questions.¹²⁶ The previous discussion demonstrates that this statement is somewhat overstated—at least some factual questions are subject to very stringent review.¹²⁷ Nevertheless, it is true that many questions of fact involve the application of judgment and discretion and that this counsels for a more deferential approach. Broadly speaking, the AB has applied a less stringent standard of review in such cases.¹²⁸ To say anything more precise requires examination of individual clauses within the covered agreements.¹²⁹ The discussion that follows provides a sketch of when panels and the AB apply deference.

Many of the most interesting problems of WTO dispute resolution require a review of the application of judgment and discretion by a member state.¹³⁰ Here the evidence is not simply a matter of objective fact, but

122. Panel Report, *United States—Transitional Safeguard Measure on Combed Cotton Yarn from Pakistan*, ¶ 7.97, WT/DS192/R (May 31, 2001) [hereinafter *US—Cotton Yarn Panel Report*].

123. *Id.*

124. See *US—Lamb* AB Report, *supra* note 106, ¶¶ 103–106.

125. See Guzman, *supra* note 47, at 12.

126. See Ehlermann & Lockhart, *supra* note 15, at 508 (stating that panels generally accept a national authority’s establishment of the facts).

127. See *supra* Part IV.A.3.

128. See Ehlermann & Lockhart, *supra* note 15, at 508.

129. Even an examination of individual covered agreements is not sufficient because the standard of review that has been applied differs from one agreement and one clause to another. *Id.* at 503.

130. *Id.* at 495.

rather one of judgment or political assessment. Given the information available to states, there may be more than one plausible conclusion. The question is to what extent should a panel or the AB either substitute its own preferred conclusion or investigate whether the member state's inference is reasonable?

Many examples of this kind of review are present in WTO jurisprudence.¹³¹ To pick just one, consider the *Argentina–Footwear* case.¹³² The AB decision in that case included the following language, referring to the standard of review adopted by the panel in examining Argentina's safeguards investigation:

[W]ith respect to its *application* of the standard of review, we do not believe that the Panel conducted a *de novo* review of the evidence, or that it substituted its analysis and judgment for that of the Argentine authorities. Rather, the Panel examined whether, as required by Article 4 of the *Agreement on Safeguards*, the Argentine authorities had considered all the relevant facts and had adequately explained how the facts supported the determinations that were made. Indeed, far from departing from its responsibility, in our view, the Panel was simply fulfilling its responsibility under Article 11 of the DSU in taking the approach it did. To determine whether the safeguard investigation and the resulting safeguard measure applied by Argentina were consistent with Article 4 of the *Agreement on Safeguards*, the Panel was obliged, by the very terms of Article 4, to assess whether the Argentine authorities had examined all the relevant facts and had provided a reasoned explanation of how the facts supported their determination.¹³³

The first portion of the inquiry asks whether Argentina had considered all relevant facts.¹³⁴ As already discussed, the appropriate review here would be quite stringent. There is no reason to believe that Argentina had special knowledge about what facts are relevant and, in any event, it is critical that the panel and the AB have a comprehensive view of the available facts.

The second portion of the inquiry concerns whether Argentine authorities "had adequately explained how the facts supported the determinations that were made."¹³⁵ The language suggests a limited inquiry: an explanation is required and, furthermore, not just any explanation will do.¹³⁶ The explanation must be "adequate."¹³⁷ The Argentine authorities had significant but not unlimited discretion. The panel is not to demand its own preferred explanation but must be satisfied with one that is merely "adequate." This relationship between the specific facts of the case and

131. See, e.g., *EC–Hormones* AB Report, *supra* note 9, ¶ 15 (discussing the standard of review as a balance of "jurisdictional competencies" both retained by member states and conceded to the WTO).

132. See generally *Argentina–Footwear* AB Report, *supra* note 118.

133. *Id.* ¶ 121.

134. *Id.*

135. *Id.*

136. See *US–Cotton Yarn* AB Report, *supra* note 96, ¶ 74 (stating that panels must "consider whether the competent authority's explanation addresses fully the nature and complexities of the data and responds to other plausible interpretations of the data").

137. *Id.*; see *Argentina–Footwear* AB Report, *supra* note 118, ¶ 121.

Argentina's decision cannot be described as a question of objective fact. It involves both the subjective evaluation of the facts (e.g., determining if there was injury or a threat of injury) and judgment about the appropriate form of response (e.g., the form of the safeguard measure).¹³⁸

Some aspects of this second inquiry are fairly objective (e.g., the increase in imports), while others are more subjective (e.g., whether the increase was sudden or whether it caused the injury).¹³⁹ The Argentine authorities, therefore, must assess the realities of the situation in Argentina and exercise their judgment. They are better positioned to consider the specifics of the situation than is a panel or the AB.¹⁴⁰ The latter are unfamiliar with the Argentine economic and social context; they have much less time to study the impact of an increase in imports, and they are poorly positioned to evaluate the extent to which safeguards are necessary to prevent or remedy serious injury.¹⁴¹

The tension between neutrality and expertise is most acute in cases involving national authorities' significant exercise of judgment.¹⁴² How should a panel or the AB review these decisions? Domestic conditions are important, and domestic authorities are much better at evaluating those conditions. Similarly, selecting methodology and drawing inferences requires expertise, and domestic authorities are more likely to have such expertise given their advantage in terms of time, resources, and contextual knowledge.

Complete deference to Argentina, however, does not seem justified. There is a great risk of bias because the protectionist tendencies of domestic governments remain.¹⁴³ Furthermore, the ability of the panels and the AB to review these decisions, although weaker than when reviewing objective evidence, remains significant. A panel or the AB can still competently review non-objective aspects of the inquiry, even if local conditions are

138. The requirement of an adequate explanation is not unique to the *Argentina–Footwear* case. It was first established in the *United States–Underwear* case, and has been repeated in several places since. See *US–Underwear* Panel Report, *supra* note 108; see, e.g., *US–Lamb* AB Report, *supra* note 106, ¶ 141; Panel Report, *Argentina–Definitive Safeguard Measure on Imports of Preserved Peaches*, ¶ 7.102, WT/DS238/R (Feb. 14, 2003); Panel Report, *Korea–Definitive Safeguard Measure on Imports of Certain Dairy Products*, ¶ 7.30, WT/DS98/R (June 21, 1999); Panel Report, *Thailand–Anti-Dumping Duties on Angles, Shapes and Sections of Iron or Non-Alloy Steel and H-Beams from Poland*, ¶ 7.256, WT/DS122/R (Sept. 28, 2000); Panel Report, *United States–Definitive Safeguard Measures on Imports of Circular Welded Carbon Quality Line Pipe from Korea*, ¶ 7.194, WT/DS202/R (Oct. 29, 2001); Panel Report, *United States–Definitive Safeguard Measures on Imports of Wheat Gluten from the European Communities*, ¶ 8.5, WT/DS166/R (July 31, 2000); Panel Report, *United States–Measures Affecting Imports of Woven Wool Shirts and Blouses from India*, ¶ 7.27, WT/DS33/R (Jan. 6, 1997); *US–Cotton Yarn* Panel Report, *supra* note 122, ¶ 7.29 (quoting *US–Underwear* Panel Report, *supra* note 108, ¶ 7.13).

139. Safeguards Agreement art. 4 (describing the application of safeguard measures).

140. See Ehlermann & Lockhart, *supra* note 15, at 501–03, 507–10.

141. *Id.*

142. *Id.* at 507–10.

143. Guzman, *supra* note 47, at 15.

relevant.¹⁴⁴

Part of the answer here is that these questions of review are simply a microcosm of the argument advanced in this article. If the review concerns legal questions of objective fact, a higher standard can be applied, and if the questions are more subjective, a lower standard should be used. But there remain questions in which objective and subjective matters are not separable. These cases call for an intermediate review. Domestic authorities must be permitted some discretion in applying their expertise and judgment, but there must nonetheless be some review to guard against abuses. This is an unavoidably imperfect process. Any discretion given to the state increases the potential for protectionist abuses, while any judicial review of subjective judgments raises the risk of error.¹⁴⁵

The task is to balance these alternative evils. In the most difficult cases, a moderate review is best suited to this task. Where neutrality is more critical and national expertise less so, there should be more review, and where the importance of these factors is reversed, there should be less.

The next question is to consider how these normative considerations compare to what panels and the AB have actually done. A thorough examination of this question would take up more space than is available here, but it is fair to say that the practice to date has largely been consistent with the approach suggested here.¹⁴⁶

For example, in the *United States–Cotton Yarn* case, the AB stated that the panel is not to prescribe a precise methodology for the collection and verification of data.¹⁴⁷ The panel allowed the United States the discretion to use its preferred method even though others were available.¹⁴⁸ One could imagine a panel inquiring as to the “best” methodology for the task at hand and requiring that the United States adopt that approach. Choosing instead to defer to American authorities’ decisions is consistent with the above argument that the domestic government will generally have greater expertise with respect to these activities. To say that a panel is not to prescribe a precise methodology, however, does not rule out a panel’s objection if the methods the domestic government used are unreasonable.¹⁴⁹ In other words, an intermediate standard is applied.

144. See Ehlermann & Lockhart, *supra* note 15, at 504–05 (noting that a panel must often confront mixed questions of law and fact which require it to balance objective and subjective factors).

145. See, e.g., Guzman, *supra* note 47, at 20–21 (“WTO review is valuable. . . as a mechanism to discourage states from engaging in conduct prohibited by WTO rules.” Yet, “WTO dispute resolution bodies . . . are more likely to err in their evaluation of the level of risk that a state’s population is prepared to accept.”).

146. See generally Ehlermann & Lockhart, *supra* note 15 (exploring how panels and the AB have approached the review of legal and factual determinations, as well as different national measures).

147. *US–Cotton Yarn* Panel Report, *supra* note 122, ¶ 7.97.

148. *Id.*

149. *Id.* (suggesting that the panel would have objected if the methods of the United States were unjustifiable).

C. Deferential Review: Science, Risk, and Risk Assessment

Finally, there is a class of cases that involve such a large degree of judgment and contextual expertise that review by a panel or the AB should be quite modest—perhaps even completely deferential.¹⁵⁰ This conclusion involves a judgment that the risk of protectionist abuse is modest and that the potential harms from a review of domestic decisions is significant.¹⁵¹ As such, it applies only in exceptional areas such as decisions about domestic risk preferences and evaluations of science.¹⁵²

The SPS Agreement requires that SPS measures (assuming they are not justified by an international standard as provided in Article 3) be based on scientific principles.¹⁵³ Article 5 of the Agreement requires, for this purpose, a risk assessment.¹⁵⁴ So, how should a panel or the AB review the decision to enact an SPS measure?

We return to the foundational claim: we must balance the advantages that the domestic government has against concerns about sham use of the SPS Agreement to achieve protectionism. The decisions involved in the implementation of an SPS measure can be disaggregated in a way that identifies some as deserving a high level of review and others as deserving extreme deference.

Consider first the evaluation of scientific data. Some data is objective and should be subject to review that is *de novo* or close to it.¹⁵⁵ Whether an international standard exists, for example, should be reviewed *de novo* as there is no meaningful concern that the panel will err.¹⁵⁶ Similarly, the panel or the AB is certainly in a position to verify that the existing studies the member state relied upon in fact exist and that the content of those studies—meaning the objective observable content, as distinct from the interpretation of the results—is as claimed. In this very narrow sense, there should be *de novo* review of scientific data.

The question of whether authorities have collected and considered all relevant evidence is much like the prior discussion of whether all evidence is considered—it is a fairly objective issue. With science, there may be a larger margin of uncertainty concerning what counts as evidence,¹⁵⁷ but otherwise it is quite similar. There should, therefore, be a high level of review on the question of whether all relevant evidence has been considered.

The important task of evaluating the data and drawing conclusions is more difficult because here there is significant scope for judgment and

150. See, e.g., Guzman, *supra* note 47, at 17-18 (stating that cases involving health and safety call for a more deferential review because "domestic governments are better at identifying the wishes of their citizens, including attitudes toward risk, and they are better at evaluating science").

151. *Id.* at 21-24.

152. See *id.* at 20.

153. SPS Agreement art. 2.2.

154. *Id.* art. 5.1.

155. OESCH, *supra* note 16, at 123.

156. See *id.* at 122-24.

157. *Id.*

application of expertise. For example, if there is no consensus of scientific opinion, should a panel or the AB require that the member state adopt the majority view or simply a credible view? In answering, one must balance at least two priorities. The first is regulatory sovereignty—states should retain the autonomy to select the level of health risk they are prepared to tolerate.¹⁵⁸ Domestic control over such decisions is important for any number of reasons, including that the willingness of individuals to accept risk may vary from state to state, as might confidence in the applicable scientific evidence.¹⁵⁹ More importantly, this decision is widely (and appropriately) accepted as one that is best left in the hands of domestic authorities.¹⁶⁰ The second priority is restricting the scope for protectionism. Full domestic control over health and safety decisions necessarily gives the states the ability to use those measures for protectionist purposes.¹⁶¹ The tension between these priorities leaves little middle ground.¹⁶² The WTO regime must, in the end, either leave policy decisions in the hands of individual states or engage in a review of those decisions, making judgments about the relevant scientific evidence and the risk tolerance of states.¹⁶³

Comparing the risks on each side, it is clear that here it is best to err on the side of deference to member states. WTO panels and the AB are quite ill equipped to engage in review of scientific judgments by member states.¹⁶⁴ Though under the rules of the DSU the panels and the AB are permitted to appoint their own experts to assist them in their deliberations,¹⁶⁵ this is a poor substitute for the skill and time required to reach reliable conclusions about the relevant science.

The potential for a panel or the AB to reach erroneous conclusions is magnified by the fact that, even after one has reached some judgment about the reliability of the relevant science, one must form a judgment

158. See Howse, *supra* note 79, at 2342.

159. *Id.*

160. *Id.*

161. See *id.* at 2354–56 (arguing that it is difficult for members to regulate themselves so that their protective measures are not more trade-restrictive than required).

162. *Id.* at 2334 (“Judgments by the WTO dispute settlement organs about what constitutes *de minimus* scientific evidence, however, would themselves entail substantive judgments of value concerning the regulatory process, begging the question of which regulatory values should determine the ‘minimum.’”); Sykes, *supra* note 71, at 355 (“WTO law must then choose between an interpretation of scientific evidence requirements that essentially eviscerates them and defers to national judgments about ‘science,’ or an interpretation that gives them real bite at the expense of the capacity of national regulators to choose the level of risk that they will tolerate.”).

163. See Theofanis Christoforou, *Settlement of Science-Based Trade Disputes in the WTO: A Critical Review of the Developing Case Law in the Face of Scientific Uncertainty*, 8 N.Y.U. ENVTL. L.J. 622, 622–23 (2000) (observing that the WTO has now “been given the binding authority to adjudicate science-based international trade disputes” and noting the new procedural and substantive issues that result).

164. *Id.* at 635 (arguing that WTO panels are not qualified to judge the plausibility of competing parties’ scientific views because they are composed of laypersons lacking in scientific knowledge).

165. DSU art. 13.

about the risk preferences of the relevant state.¹⁶⁶ Is the member state willing to accept some increase in, say, the risk to human health in exchange for a product that is lower cost? This question arose in both the *EC–Hormones* and the *EC–GMOs* cases.¹⁶⁷ The risk tolerance of a state is, of course, the product of complex political forces, and it is hard to imagine a panel reaching a reliable conclusion on the subject.¹⁶⁸ Here, the domestic government has much greater expertise, and so the case for deference is stronger.

Turning to what is actually done in WTO cases, the AB appears to have established a higher standard of review than the above discussion recommends.¹⁶⁹ In reviewing the scientific evidence used in risk assessments under the SPS Agreement, the AB has indicated that panels are to evaluate the scientific evidence with little deference to the views of the relevant domestic government.¹⁷⁰ For example, in the *Japan–Apples* case, the AB stated that “Japan’s submission that the Panel was obliged to favour Japan’s approach to risk and scientific evidence over the views of the experts conflicts with the Appellate Body’s articulation of the standard of ‘objective assessment of the facts.’”¹⁷¹ A similar attitude was evidenced in the *Australia–Salmon* case, where the AB stated that panels “are not required to accord to factual evidence of the parties the same meaning and weight as do the parties.”¹⁷² Matthias Oesch goes so far as to conclude that the “panels’ approach has overall amounted to a *de novo* standard of review with respect to the relevance and adequacy of scientific data.”¹⁷³ This conclusion may overstate the extent to which panels review states’ scientific judgments, in particular because states are entitled to use minority scientific views in forming their judgments.¹⁷⁴ The lack of deference to state decisions must be viewed within the context of a set of substantive rules that give states considerable flexibility in how they evaluate science. Nevertheless, it is accurate to say that the evaluation of science is subject to a significant review.

In contrast to the treatment of scientific evidence, panels are to give considerable deference to “risk management.”¹⁷⁵ “The determination of

166. Howse, *supra* note 79, at 2342 (“[S]cience cannot tell us just how conservative or protective it is reasonable to be in the presence of a given level of error or uncertainty in a scientific assessment of risk. In a democracy, this will depend on citizens’ preferences about risk.”).

167. See generally *EC–Hormones* AB Report, *supra* note 9; Panel Report, *European Communities–Measures Affecting the Approval and Marketing of Biotech Products*, WT/DS291/R, WT/DS292/R, WT/DS293/R (Sept. 29, 2006).

168. See OESCH, *supra* note 16, at 141–42.

169. See *infra* notes 171–172 and accompanying text.

170. See *infra* notes 171–172 and accompanying text.

171. Appellate Body Report, *Japan–Measures Affecting the Importation of Apples*, ¶ 165, WT/DS245/AB/R (Nov. 26, 2003).

172. Appellate Body Report, *Australia–Measures Affecting Importation of Salmon*, ¶ 267, WT/DS18/AB/R (Oct. 20, 1998) [hereinafter *Australia–Salmon* AB Report].

173. OESCH, *supra* note 16, at 123.

174. *EC–Hormones* AB Report, *supra* note 9, ¶ 194.

175. See *Australia–Salmon* AB Report, *supra* note 172, ¶¶ 198–200; Panel Report, *European Communities–Measures Concerning Meat and Meat Products (Hormones)*, Com-

the appropriate level of protection . . . is a *prerogative* of the Member concerned and not of a panel or of the Appellate Body.”¹⁷⁶ The justification for deference in this context is that risk management is thought to be non-scientific.¹⁷⁷ This is consistent with this article’s general argument in the sense that risk management is something that requires a great deal of local knowledge. Domestic governments are therefore better equipped to carry it out than are panels or the AB.

Therefore, the existing rules governing the standard of review in SPS cases leave judgments about risk tolerance to the states while allowing WTO adjudicatory bodies to review the merits of the relevant science. There are at least two problems with this approach. First, with respect to evaluating science, domestic authorities have a significant advantage over the WTO’s adjudicatory organs.¹⁷⁸ Scientific evidence is often disputed within the scientific community, making it necessary for policymakers to judge the relative merits of competing scientific claims. This is difficult enough for government officials working with the advantage of time, resources, and the freedom to access expertise in any suitable way. For a panel or the AB the task is much more difficult. Panelists and AB members are not chosen for their scientific expertise and they face time constraints as well as limited resources, staffing, and budgets for their inquiry.¹⁷⁹ There is no realistic prospect that they could carry out a thorough and informed evaluation of scientific evidence.¹⁸⁰

Second, a state’s evaluation of science is itself influenced by domestic priorities and cultural values.¹⁸¹ The use and evaluation of scientific evidence includes a significant subjective component.¹⁸² States may, in good faith, disagree on the meaning and importance of scientific evidence, and these differences may reflect country-specific attitudes toward scientific proof and inference.¹⁸³ Understandably, domestic governments have a better sense of local attitudes toward science than does a panel or the AB.

The difference between the approach laid out in AB jurisprudence and the approach advocated in this article turns primarily on the relative abilities of states and panels to evaluate scientific evidence. If panels could objectively and accurately evaluate that evidence, and if that evaluation

plaint by Canada, ¶¶ 8.100, 8.163, WT/DS48/R/CAN (Aug. 18, 1997) [hereinafter *EC–Hormones, Complaint by Canada*] (defining “risk management”). *But see EC–Hormones AB Report*, *supra* note 9, ¶ 181 (rejecting the panel’s distinction between “risk assessment” and “risk management”).

176. *Australia–Salmon AB Report*, *supra* note 172, ¶ 199.

177. *EC–Hormones, Complaint by Canada*, *supra* note 175, ¶¶ 8.100, 8.163.

178. *See Guzman*, *supra* note 47, at 20–23, 38.

179. *Id.* at 23; *see* DSU art. 8; *see also* Christoforou, *supra* note 163, at 647–48.

180. It is true that panels and the AB can consult with experts to assist them in their decision. DSU art. 13. Although this improves their capacity to handle scientific issues, it is far from a substitute for the capabilities of domestic authorities.

181. *See Guzman*, *supra* note 47, at 22; *see also* Steve Charnovitz, *Improving the Agreement on Sanitary and Phytosanitary Standards*, in *TRADE, ENVIRONMENT, AND THE MILLENNIUM* 171, 171–72 (Gary P. Sampson & W. Bradnee Chambers eds., 1999).

182. *See Charnovitz*, *supra* note 181, at 171–72; *Guzman*, *supra* note 47, at 21–24.

183. *See Guzman*, *supra* note 47, at 21.

required no particularized knowledge regarding the priorities and attitudes of individual states, then the existing approach would be defensible. As already discussed, however, this is an unrealistic picture of how disputes about science play out. Because domestic governments have a greater ability to evaluate scientific evidence than do panels, the review of government decisions should be quite deferential.

There is, of course, a risk that a state will use the SPS rules as a justification for protectionism, even if it has no legitimate concern about health and safety.¹⁸⁴ This concern, however, does not undermine the argument for deference. First, given that risk management is left to the states and is subject to quite deferential review, a stringent review of the science used adds no more than modest discipline to the system. Certainly, as long as there is agreement that a product poses some degree of risk, a state can always opt for an exposure to that risk that is sufficiently low as to achieve their trade objectives.¹⁸⁵ If there is no scientific consensus regarding the presence or absence of risk, then the question becomes one of whether a product is sufficiently safe.¹⁸⁶ This discussion must necessarily include consideration of a state's willingness to expose itself to an uncertain risk—which returns the analysis to the risk management question.¹⁸⁷ To date, the AB has allowed panels to review this secondary risk management question (exposure to uncertain risks), but it is not clear why secondary risk management is treated differently from the primary risk management decision (exposure to known risks).¹⁸⁸

Second, although the potential for abuse exists, it is no greater under the SPS Agreement than it is in other trade disputes. What has changed is the competence of panels and the AB.¹⁸⁹ Because these bodies are poorly situated to assess science or risk tolerance, the balance between domestic expertise and WTO neutrality tilts further toward domestic expertise.¹⁹⁰ With respect to both the evaluation of science and the risk tolerance of member states, it is sensible to leave almost complete discretion with

184. See *id.* at 15; see also SPS Agreement art. 5.8.

185. See SPS Agreement Annex A ¶ 5 (defining the appropriate level of protection as that “deemed appropriate by the Member”).

186. See *id.* art. 3.3 n.2.

187. See *supra* notes 175–177 and accompanying text.

188. See *Australia–Salmon* AB Report, *supra* note 172, ¶¶ 198–200; *EC–Hormones*, *Complaint by Canada*, *supra* note 175, ¶¶ 8.100, 8.163 (defining “risk management”). But see *EC–Hormones* AB Report, *supra* note 9, ¶ 181 (rejecting the panel’s distinction between “risk assessment” and “risk management”).

189. See Christoforou, *supra* note 163, at 635 (noting that WTO panels lack scientific knowledge and thus are not qualified to judge competing parties’ scientific views).

190. See Guzman, *supra* note 47, at 17–18, 23 (arguing that panels and the AB are poorly equipped to evaluate scientific evidence, while domestic governments are better at evaluating science). It is also worth noting that deference may sometimes be required to preserve the legitimacy and stability of the dispute settlement system. When a review at the WTO impinges too much on traditional areas of sovereignty, even if judgments require only limited domestic expertise, it may be appropriate to defer to national authorities. See Ehlermann & Lockhart, *supra* note 15, at 493–94 (noting that panels and the AB are criticized for their level of scrutiny of national measures and that they are accused of usurping WTO members’ powers).

domestic authorities. To do otherwise is to ask panels and the AB to form judgments in areas in which they are supremely unqualified.

Finally, the analysis must consider those requirements within the SPS Agreement that do not require judgments about science.¹⁹¹ Consider, in particular, the set of requirements that encourage transparency.¹⁹² These include an obligation to publish regulations promptly,¹⁹³ to give producers time to adjust before a measure comes into effect,¹⁹⁴ to establish “enquiry points” to answer questions relating to the measure,¹⁹⁵ and to give other member states an opportunity to comment and discuss the proposed measure.¹⁹⁶ Additionally, the risk assessment itself must be transparent.¹⁹⁷ Members must answer all reasonable questions regarding the procedures used in the assessment, including the factors taken into consideration and the “determination of the appropriate level of sanitary or phytosanitary protection.”¹⁹⁸ The requirements impose a significant degree of openness on the adoption of an SPS measure and should be subject to a high level of review. These are objective questions of fact and a panel or the AB can review them accurately. The enforcement of transparency requirements also serves to guard against protectionist abuse as it makes it more difficult for a member state to assert a pretextual justification.¹⁹⁹ “Measures that comply with the procedural requirements of the SPS Agreement but that look to the rest of the world like protectionist measures will be more costly to adopt because affected states will put political pressure on the enacting member.”²⁰⁰

D. The Special Case of the AD Agreement

As previously mentioned, the Anti-Dumping Agreement provides its own, deferential standard of review. In particular, the panel is to accept a member state’s views as long as it determines that the state’s establishment of the facts was “proper” and that the state’s evaluation of those facts was “unbiased and objective.”²⁰¹ This deferential standard of review provides states with considerable leeway in establishing the facts necessary to pursue their preferred policies. Notice that the standard of review focuses entirely on process rather than substance.²⁰² A panel can refuse to accept

191. See, e.g., SPS Agreement arts. 7-9.

192. *Id.* art. 7.

193. *Id.* art. 7, Annex B ¶ 1.

194. *Id.* Annex B ¶ 2.

195. *Id.* Annex B ¶ 3.

196. *Id.* Annex B ¶ 5(d).

197. *Id.* art. 5.1, Annex B ¶ 3(c).

198. *Id.* Annex B ¶ 3(c).

199. Guzman, *supra* note 47, at 35.

200. *Id.*; see Howse, *supra* note 79, at 2330 (arguing that the SPS “provisions can be, and should be, understood not as usurping legitimate democratic choices for stricter regulations, but as enhancing the quality of rational democratic deliberation about risk and its control”).

201. AD Agreement art. 17.6(i).

202. Croley & Jackson, *supra* note 23, at 208 n.65 (noting that Article 17.6 gives the panels an obligation to evaluate the fact-finding process).

a member state's view of the facts only if it finds the process was defective—either because the establishment of facts was not proper or because the evaluation of the facts was biased.²⁰³ Although states must avoid providing biased evidence, there is little to prevent them from shading the facts to suit their needs.

With respect to legal interpretations, the panel “shall interpret the relative provisions of the Agreement in accordance with customary rules of interpretation of public international law.”²⁰⁴ This requirement refers to the interpretive strategies set forth in Articles 31 and 32 of the Vienna Convention on the Law of Treaties.²⁰⁵ If the panel finds that “a relevant provision of the Agreement admits of more than one permissible interpretation, the panel shall find the authorities’ measure to be in conformity with the Agreement if it rests upon one of those permissible interpretations.”²⁰⁶ In other words, if a member state can justify its measure under a permissible interpretation—even if other permissible interpretations exist—then the panel must conclude that the measure is consistent with the AD Agreement.²⁰⁷

The implication of the above standard of review for legal questions is to greatly expand the permissible set of anti-dumping measures that a state may undertake. Different states can, for example, rely on mutually inconsistent interpretations of the agreement to justify their respective measures.²⁰⁸ As long as both interpretations are “permissible,” a panel or the AB will find both interpretations sufficient to permit the measures.²⁰⁹

The significant deference to the member states enacting an anti-dumping measure is arguably consistent with the substantive provisions in the AD Agreement. These provisions require that the enacting state carry out an investigation prior to adopting anti-dumping measures,²¹⁰ provide details about the gathering of evidence and notice to interested parties,²¹¹ and so on.²¹² Absent Article 17.6, however, it is easy to imagine a panel’s more substantive review of both facts and law.²¹³ The fact that national authorities have conducted an investigation does not imply that the investi-

203. AD Agreement art. 17.6.

204. *Id.* art. 17.6(ii).

205. *Hot-Rolled Steel* AB Report, *supra* note 34, ¶ 57; Daniel K. Tarullo, *The Hidden Costs of International Dispute Settlement: WTO Review of Domestic Anti-Dumping Decisions*, 34 *LAW & POL’Y INT’L BUS.* 109, 119–20 (2002); see Vienna Convention arts. 31–32.

206. AD Agreement art. 17.6(ii).

207. See *id.*

208. See Croley & Jackson, *supra* note 23, at 200 (arguing that if the national interpretation is within the set of “permissible” interpretations, “the panel must defer to the interpretation given the provision by the national government”).

209. *Id.*

210. See AD Agreement art. 5.

211. *Id.* art. 6.

212. See Ehlermann & Lockhart, *supra* note 15, at 507 (arguing that the AD Agreement “gives a pre-eminent position to the national investigation”).

213. *Id.* (“The role of the panel is . . . confined to reviewing the investigation and determination that has been carried out at the national level.”).

gation reached the correct conclusion or that it properly interpreted relevant WTO law.

Article 17.6 of the AD Agreement reflects the decision of the member states that panels should refrain from aggressive review of anti-dumping measures.²¹⁴ The approach that the AD Agreement requires is inconsistent with this article's normative proposal. It is perhaps telling that the drafters of the AD Agreement felt the need to include a specific standard of review provision.²¹⁵ Without that provision, it is difficult to see how a panel would have concluded that such a deferential review of AD measures would be appropriate.

Conclusion

To date the WTO has failed to provide clear and predictable principles to govern the standard of review. As a result, member states and panels are unable to accurately predict the outcome of individual cases. This article offers a simple yet principled approach to the standard of review problem that can be applied to the full range of disputes at the WTO. At the heart of this approach is a recognition that domestic governments and panels have distinct strengths and weaknesses. States have information about domestic priorities and preferences, allowing them to adjust domestic policy to serve their country's needs. Because they are responsive to domestic needs, however, member states cannot be relied upon to evaluate either facts or legal rules in an unbiased fashion.

Panels and the AB, however, are poorly positioned to assess domestic preferences and priorities. They are not designed to perform this role, and they lack the resources and skills to carry it out. The key asset of the panels and the AB is their neutrality. Unlike member states, they have no interest in serving the interests of one member over those of another. Panels and the AB are best suited for the task of providing unbiased interpretations of WTO agreements where required.²¹⁶

This simple tradeoff between expertise and neutrality can and should guide the determination of appropriate standards of review. On the one hand, where it is possible to evaluate relevant facts using objective criteria and without reference to domestic priorities, panels and the AB should apply a high level of review. Where, on the other hand, there is little risk of bias and a decision requires a great deal of local knowledge, there should be deference to domestic policymakers. In the large number of cases that lie between these two extremes, the standard of review should reflect the relative importance of neutrality and domestic expertise.

Existing WTO jurisprudence is generally in line with the proposed approach.²¹⁷ The one area where actual practice diverges significantly from the normative discussion is in the area of risk and risk assessment.

214. See AD Agreement art. 17.6.

215. *Id.*

216. See *supra* Part III.B.

217. See *supra* Part IV.

This area is one in which domestic judgments and priorities, along with the other advantages domestic governments have in evaluating science, make domestic governments much better decisionmakers than panels. Although there are concerns about neutrality in this area, the importance of localized expertise outweighs those concerns. For these reasons, the appropriate standard of review in both risk management and risk assessment is one of great deference. To date the WTO has shown deference in the area of risk management, but not in the area of risk assessment.²¹⁸ This may reflect a difference in attitudes about the extent to which scientific judgments can be reviewed objectively by a panel without detailed knowledge of local attitudes toward science.

Though the difference between the proposal advanced in this article and actual practice is significant, the larger lesson is that the standard of review has been appropriately determined in most areas. What is missing is an explicit embrace of the tradeoff between neutrality and expertise. If the AB were to adopt this as a general approach to the problem, everyone involved would have a much easier time anticipating the standard of review that would be applied in a given case, and panels and the AB would have clear guidelines to determine the appropriate standard when faced with new situations.

218. See *Australia–Salmon* AB Report, *supra* note 172, ¶¶ 198–200; *EC–Hormones*, Complaint by Canada, *supra* note 175, ¶¶ 8.100, 8.163 (defining “risk management”). But see *EC–Hormones* AB Report, *supra* note 9, ¶ 181 (rejecting the panel’s distinction between “risk assessment” and “risk management”).