



European
University
Institute

ROBERT
SCHUMAN
CENTRE FOR
ADVANCED
STUDIES

WORKING PAPERS

RSCAS 2016/31

Robert Schuman Centre for Advanced Studies
Global Governance Programme-220

The Gang that Couldn't Shoot Straight
The Not So Magnificent Seven of the WTO Appellate Body

Petros C. Mavroidis

European University Institute
Robert Schuman Centre for Advanced Studies
Global Governance Programme

The Gang that Couldn't Shoot Straight
The Not So Magnificent Seven of the WTO Appellate Body

Petros C. Mavroidis

EUI Working Paper **RSCAS** 2016/31

This text may be downloaded only for personal research purposes. Additional reproduction for other purposes, whether in hard copies or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper, or other series, the year and the publisher.

ISSN 1028-3625

© Petros C. Mavroidis, 2016

Printed in Italy, June 2016

European University Institute

Badia Fiesolana

I – 50014 San Domenico di Fiesole (FI)

Italy

www.eui.eu/RSCAS/Publications/

www.eui.eu

cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), created in 1992 and directed by Professor Brigid Laffan, aims to develop inter-disciplinary and comparative research on the major issues facing the process of European integration, European societies and Europe's place in 21st century global politics.

The Centre is home to a large post-doctoral programme and hosts major research programmes, projects and data sets, in addition to a range of working groups and *ad hoc* initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration, the expanding membership of the European Union, developments in Europe's neighbourhood and the wider world.

Details of the research of the Centre can be found on:

<http://www.eui.eu/RSCAS/Research/>

Research publications take the form of Working Papers, Policy Papers, and e-books. Most of these are also available on the RSCAS website:

<http://www.eui.eu/RSCAS/Publications/>

The EUI and the RSCAS are not responsible for the opinions expressed by the author(s).

The Global Governance Programme at the EUI

The Global Governance Programme is one of the flagship programmes of the Robert Schuman Centre for Advanced Studies at the European University Institute (EUI). It aims to: build a community of outstanding professors and scholars, produce high quality research and, engage with the world of practice through policy dialogue. At the Global Governance Programme, established and early career scholars research, write on and discuss, within and beyond academia, issues of global governance, focussing on four broad and interdisciplinary areas: European, Transnational and Global Governance; Global Economics; Europe in the World; and Cultural Pluralism.

The Programme also aims to contribute to the fostering of present and future generations of policy and decision makers through its unique executive training programme, the Academy of Global Governance, where theory and "real world" experience meet. At the Academy, executives, policy makers, diplomats, officials, private sector professionals and academics, have the opportunity to meet, share views and debate with leading academics, top-level officials, heads of international organisations and senior executives, on topical issues relating to governance.

For more information: <http://globalgovernanceprogramme.eui.eu>

Abstract

The WTO Appellate Body (AB) has produced a volume-wise important body of case law, which is often difficult to penetrate, never mind classify. Howse (2016) has attempted a very lucid taxonomy of the case law using the standard of review as benchmark for it. His conclusion is that the AB is quite cautious when facing nondiscriminatory measures, especially measures relating to the protection of human life and health, while it has adopted a more intrusive (into national sovereignty) standard when dealing with trade measures (like antidumping), which are by definition discriminatory as they concern imports only. In my response, I share his basic conclusion with no buts and ifs. I simply add that this approach is not the outcome of a process that mandates this standard of review, but simply a political (e.g., nonlegal) reaction aimed at placating its clientele, the WTO membership.

Keywords

WTO; Dispute Settlement; Appellate Body

JEL Classification: K40

1. The Argument*

This paper is a reaction to Howse (2016). The brass tracks of Howse's argument is that the WTO Appellate Body (AB) has adopted a rather deferential standard towards domestic policies, which have to be nondiscriminatory anyway, and a more intrusive attitude when dealing with trade instruments, which by definition concern imports only. I propose to entertain my response in three parts.

In the brief Section 2, I will ask the question whether there are statutory underpinnings supporting the approach privileged by the AB. My response is no. If at all, the framers of the WTO wanted panels and the AB to adopt a deferential standard of review, when dealing with antidumping, and not with domestic instruments. The AB undid the statutory premise, and applied deference not to disputes regarding trade-, but to disputes regarding domestic instruments (policies). This is all judge made law, it reflects the quintessential belief of the AB how things should be done, and Howse is absolutely right to single out this element.

In Section 3, I move to discuss the practice of the AB decision. I will place it in its historical context, and explain briefly the GATT case law first, which was toing and froing between deference and intrusiveness when dealing with domestic instruments. The AB in its remarkable US-Shrimp report decided at least one contentious issue for good, namely that domestic instruments would not be judged illegal simply because they had been unilaterally adopted. The AB has not nevertheless, adopted a consistent standard of review across domestic instruments, and has outlawed many of them on shaky evidence. Still, it will not tinker with measures aiming to protect human health. That much is for sure. The AB on the other hand, when dealing with border instruments, has imposed stringent requirements on key disciplines, like causality. The AB however, has sugarcoated this approach through soft remedies, and generous (for the regulator) understanding of the obligations regarding continued imposition of duties (sunset reviews).

In Section 4, I will advance my explanation why the AB has adopted this approach. In my view, the approach is warranted and rests on solid grounds. Parties have little incentive to reveal the rationale behind adoption of their domestic policies, and, absent similar knowledge, it is often hard to decide on who is right or wrong. Risk averse courts, worried about the institutional implications of false positives, might rationally prefer to avoid outlawing challenged measures. Alas, the AB has not followed a similar path in developing its approach. It has behaved as a political body reacting to signs of times, a 'Warren Court' of trade, rather than a court that has developed its own understanding of the rationale for, and the objective function of the agreements that it is called to interpret. The quest for contextual understanding of the key terms remains, as a result, elusive. Predictions regarding the manner in which it will treat similar cases in the future are equally impossible to make.

Section 5 recaps my main conclusions.

2. Standard of Review: Statutory Language and Practice

The AB did not start from a clean slate when developing its standard of review. In fact, it was called to use one generic standard of review across all cases, and another, arguably more deferential, in antidumping disputes. The AB developed its standards of review against this explicit will of the framers to treat a trade instrument, antidumping, in more deferential manner than any other measure coming under the ambit of the WTO.

* For helpful discussions and comments on previous drafts, I would like to thank Bill Davey, Henrik Horn, and Damien J. Neven for many discussions of the issue and comments on previous drafts.

2.1 Two Statutory Standards of Review

Article 11 of the DSU (Dispute Settlement Understanding) requests from panels to make an ‘objective assessment’ of the matter before them, without any further clarifications. Of interest to our discussion is not the substantive content of this standard. What we care about is the fact that the DSU provides for one standard to be applied across all cases, irrespective whether we deal with trade or domestic instruments (policies). There is only exceptional standard, that reflected in Article 17.6(ii) of the Agreement on Antidumping (AD), which was meant to apply only to disputes regarding impositions of AD duties.

This provision (Article 17.6(ii)) repeats in the first sentence a constraint that the DSU also acknowledges (Article 3.2), namely, that panels (and the AB) should have recourse to the customary laws of public international law, when interpreting the WTO contract. The second sentence though, calls for panels (and the AB) to refrain from going any further when they have encountered an interpretation of the terms that is ‘permissible’. The working hypothesis for the framers must have been that more than one permissible interpretation was possible, at least on occasion. It is clear, that the US, the instigator of this provision, had in mind a deferential standard of review.¹ Various negotiating documents support this conclusion.²

2.2 The AB, Master and Commander

Against this background, the AB performed two innovations: first, it merged the two standards into one, and, second, it elevated protection of human health to the pedestal of global values.

2.2.1 And the Two Shall Become One

With two early exceptions,³ panels have sided with the AB report in US-Hot Rolled Steel (§62), where the AB saw no disharmony between the two statutory standards of review discussed above. De facto, this meant the end of deference in antidumping disputes.

2.2.2 The Importance of Objective Sought

In EC-Asbestos, the AB underscored that it would be shaping its standard of review in light of the importance of the regulatory objective pursued. In subsequent case law, it became clear what the AB meant: a deferential approach when protection of human health was at stake, and a more intrusive standard when other societal preferences were at stake.⁴

3. From Broad Lines to Deets

So far I have attempted to explain the broad lines of the AB standards of review? What does this all mean in practice? How has the AB operationalized its preferred standard of review?

3.1 Domestic Instruments: Belts and Suspenders (in Principle)

The AB inherited GATT case law, which in part it kept, and in part it rejected. GATT panels had struggled with the question of standard of review. Although formally they had adopted a “no effects

¹ Croley and Jackson (1996).

² Stewart et al. (1993) refer to many of them.

³ Both are discussed in Mavroidis (2016), Chapter 2, volume 2.

⁴ Mavroidis (2016) Chapter 7, volume 1, and Sykes (2003) discuss this case law.

cum no intent” standard, which, prima facie might sound quite intrusive (as it allows panels to outlaw measures that might have no demonstrable protectionist effect, while not inquiring into the intent at all), they have in fact attempted to show deference towards regulatory intent, to the extent that they were persuaded that it was not protectionist. We explain.

3.1.1 Debates in GATT

GATT panels struggled with the legal discipline regarding domestic instruments. The originally prevailing view was that the rationale for intervention should matter only as grounds justifying deviation from an obligation. Fearing that this approach would lead to an understanding of the GATT as instrument for deregulation (rather than nondiscrimination, its original purpose), two reports endorsed the “aims and effect”-test, where likeness of goods would be function of the regulatory purpose. Measures aiming to protect noneconomic preferences would thus, be judged GATT-consistent. GATT panels did not elaborate a meticulous intent-test, when doing so. They would satisfy themselves on absence of protectionism on little evidence to this effect.⁵

Case law was thus oscillating between an analysis where regulatory intent was irrelevant, and where it was the centerpiece of the inquiry.

There are two reasons why deference should not be regarded as the dominant standard of review. First, the quintessential report reflecting the deferential standard (US-Taxes on Automobiles) remained un-adopted, and consequently, of limited legal value. Second, because in a subsequent case, US-Tuna I (Mexico), the panel outlawed a US measure simply because it was unilateral, without even reaching the stage to ask the question regarding the relevance of the aim. This panel did not even ask itself how could it ever be right in light of the fact that all measures challenged under Articles III and/or XI of GATT were unilateral, and here it was dealing with a defense against a violation of the latter provision.

At the very least thus, the GATT did not leave a legacy of consistency regarding the treatment of domestic policies. The AB inherited this legacy. How did it go about it?

3.1.2 First, a Huge Sigh of Relief

US-Shrimp is a remarkable report. The AB set the record straight when holding that unilateral policies are not inconsistent simply because they are unilateral. It was a head on reversal of the US-Shrimp I (Mexico) GATT case law, which had stated that the opposite was true. But of course the GATT panel could not have been right. The GATT was a tariff bargain supported by a negative integration regime. Its members were not required to adopt common policies. They could design their domestic policies to their liking as long as they applied them in nondiscriminatory manner.

Nondiscrimination was legalese for absence of protectionism, an elusive concept, which the GATT-think reduced to a negotiable instrument, tariffs. All accounts of the GATT that extensively refer to the negotiating record, from early on Wilcox (1949), to the seminal work of Baldwin (1970), and Jackson (1969), to most recent inquiries like that by Irwin et al. (2008) agree on this score. Under the circumstances, US-Tuna I (Mexico) could have never been right, and US-Shrimp was a welcome sea change.

Of course, it could never be the case that unilateral policies are necessarily legal either. A discrimination (protectionism) test had to be devised as well. The question for the AB was whether it should pick one of the two tests already developed in GATT case law, or whether it should develop its

⁵ Hudec (1998) offers an excellent description of the “aims and effect” test, and why inquiries into regulatory intent, irrespective of the manner in which the test had been applied in case law, was necessary for GATT to be understood as an instrument for nondiscrimination, and nothing beyond that.

own. In name, it adopted the “marketplace”-test. Regulatory intent would be of relevance in a very limited set of circumstances when the burden of proof rests with the complainant (Chile-Alcoholic Beverages), and quite relevant when the burden of proof would shift to the defendant that would have to have recourse to the list of general exceptions (Article XX of GATT) in order to justify violation of an obligation. In practice, it managed to create a mess through a mix of contradictions, open-ended statements, and difficult to grasp decisions, as I will try to explain in what now follows.

3.1.3 Then, the AB Missed the Compass

I think I am not the only professor of WTO law that, over the years, has found it exceedingly hard to explain the nondiscrimination test devised by the AB.

In Japan-Alcoholic Beverages II, consumers would define likeness, preferably through the use of the best predictor we have, cross-price elasticity. Less favourable treatment, at least in the panel’s (clearer) reading of the situation, did not exist because the measure imposed a higher burden on imported competing goods, and the defendant had not advanced any policy rationale to justify its choices. Indeed, as the panel observed, Japan only *ex post facto* came up with an explanation that differential taxation was meant to protect traditional production processes.

Every report that followed instead of improving the situation, substantially muddied the waters. In Korea-Alcoholic Beverages, recourse to econometric indicators and noneconometric indicators was put at par. This could well be the case, if the former yield no adequate response (because, for example, a good has been *de facto* banned from a market through discriminatory domestic taxation). Surprisingly, price of goods was absent from the list of relevant factors, as if purchasing decision for the majority of consumers are not function of scarcity of monetary resources.

Likeness was function of competitive relationship and sharing the same six-digit classification (Japan-Alcoholic Beverages II). The AB added that the same would be the case when goods are in very intensive competitive relationship. Then comes Philippines-Distilled Spirits. The AB is dealing with goods that do not share the same six-digit classification, but which, in its view are still like, because they are in intense competitive relationship. It refers to studies that measured the cross-price elasticity coefficient, and the resulting range was between 0.01 and 0.07. This is quite low by any reasonable benchmark. A value of 0.01 would imply that a tax on imports that increases the price of imports by 50 percent would increase the volume for the domestic product by 0.5 percent, which is close to nothing. The products, hence, should have been considered almost independent goods and not market-like, as far as degree of substitution goes. If elasticity was only marginally smaller, and it equaled 0, the two products would be completely independent. The AB tries to make up for this arguing that the goods share the same end uses etc. A bicycle and an airplane though share the same end use (transport), but no one would call them like goods.

Alas, there is worse. In EC-Asbestos, likeness is delinked from market analysis. Two goods are like if a ‘reasonable consumer’ so thinks, and no market evidence to this effect is warranted. The judgment of the members of the AB (the ‘reasonable consumer’, for all practical purposes) can substitute for that. This was bad news for Canadians, who in the eyes of the AB must be unreasonable.⁶ The good news for the defendant was that, because the two pairs of construction material (with-, and without asbestos) were considered unlike, it did not have to defend its policies any further. Likeness was function of the properties of the goods concerned, and their potential negative impact on human

⁶ To avoid misunderstandings, in Grossman et al. (2013), I published my views when stating that EU should have prevailed but not because the two goods were unlike. It should have prevailed because France was not affording less favourable treatment when banning sales of health-impairing goods. It is the privilege of any WTO member to define its level of protection, and France can have a higher level of protection than Canada if it so wishes, but those who disagree are not necessarily unreasonable.

health. This case law established the importance of the objective as relevant criterion to decide on the degree of deference that panels should show to the regulator.

And when we thought we had seen the end of innovation, along came the AB report on Argentina-Financial Services, where likeness was presumed this time. The AB invented a concept that it had never used before, and made us wonder what exactly it meant with “presumption of likeness”.⁷

The interpretation of less favourable treatment presents us with similar discomfort. In Dominican Republic-Import and Sale of Cigarettes, following toing and froing towards every imaginable direction, the AB seemed to have set the record straight when stating that measures creating disparate effects would not be judged inconsistent, if the rationale for the adoption of measure was unrelated to the origin of the good. Then comes EC-Seal Products. The AB deplors our poor reading of the unambiguous and hard to misunderstand paragraphs in its previous report, and states that presence of disparate effects means less favourable treatment.

This case law casts doubt on the generic relevance of the deference standard to domestic policies. Deference was effectively limited to cases where human health is at stake (EC-Asbestos), and not when other societal preferences are advanced.⁸ In fact, a number of domestic policies have been outlawed on shaky evidence.

Korea-Various Measures on Beef offers an adequate illustration of this point. The AB outlawed a nondiscriminatory measure (dual retailing) that, in its own admission, was genuinely aiming at protecting the stated regulatory objective. It held that the measure was unnecessary, but did not provide any evidence regarding the disparate trade impact of the measure. If the intent was legitimate, and disparate trade effects were absent,⁹ on what evidence was the measure judged GATT-inconsistent?

The situation is alas, worse when it comes to disputes under the TBT- and the SPS Agreements. In EC-Hormones, the AB accepts the zero-risk policy (whatever it means) practiced by the European Union, and goes so far as to suggest that the precautionary principle is not confined to Article 5.7 of SPS (as some, including me, might have thought), but permeates the whole agreement.¹⁰ A few years later, Japan invokes zero-risk policy with respect to fire blight, an apples disease. In Japan-Apples, the AB finds nothing wrong with the risk assessment supplied by the defendant, but outlaws the measure because it considers it unnecessary. This report looks like an one bad apple, as Neven and Weiler (2006) have suggested, and lends suspicion to those arguing that the AB serves one sauce for the goose and one for the gander. The probably more likely explanation though is that the objective pursued by the EU (protection of human health) defined endogenously the standard of review for the AB.

3.1.4 Deference Yes, but to a Limited Class of Domestic Policies, Namely ...

So what do I conclude from the above? First, the AB cannot be accused of having provided a clear methodology that will enable it (and subsequent panels) to distinguish wheat from chaff. There is no nondiscrimination test, the proxy to decide on protectionist behaviour. Some early honourable attempts have been followed by subsequent confusion as to what the original attempts amounted to, and, even more to the point, where we now stand on this score. There are four key terms in the body of

⁷ Of course, likeness can never be presumed, since different consumers in different markets may react in different ways to the same pair of goods, and, absent some market research, the competitive relationship between the goods cannot be established.

⁸ Sykes (2003).

⁹ The panel had found that Korea was routinely absorbing the legal import quota in place. This was quite normal, in light of the difference between the world- and the Korean price for beef.

¹⁰ Davey (2006) offers a very comprehensive analysis.

Article III of GATT: “like”; directly competitive or substitutable”; “in excess”; “applied so as to afford protection”. There is confusion regarding the understanding of three of them. Only the term “in excess” has been consistently understood as denoting arithmetical difference.

Second, that the AB has adopted a deferential standard of review but towards some domestic policies, namely measures aiming to protect human life and/or health.

Third, that it did not provide any reason for doing so other than the fact that in its view risks to human life and/or health deserve maximum deference.

This leads me to conclude that the deferential standard towards measures aiming to protect public health is not the outcome of analysis that led the AB to take the view that it was facing serious evidentiary issues when dealing with challenges against domestic policies. It is a knee-jerk reaction to what is to most the highest value, the quintessential human right.

3.2 Trade Instruments: Gung Ho (in Principle)

The AB has inherited a very rich case law in the realm of trade instruments from the GATT-era. Although disputes regarding tariff treatment were infrequent in the GATT, case law in antidumping provided a lot of ammunition for heated discussion across the membership.

3.2.1 GATT Debates

Disputes regarding trade instruments multiplied following the advent of the Tokyo round codes. Case law especially in antidumping and countervailing stands for the proposition that remedies in this context should be retroactive. If Home was found to have imposed antidumping duties illegally so, Foreign could request repayment of all duties perceived between the original imposition and the time when the panel had pronounced on the illegality. This remedy did not go down well, especially with the EU and the US. In fact, as Hudec (1993) explains, the introduction of retroactive remedies was a reason why the rate of adoption of GATT panel reports fell in the eighties.

In Mavroidis (2016a), I argued that the EU and the US attempted to reverse this trend during the Uruguay round negotiations, to no avail. Law (the DSU) did not formally adopt their motion to introduce prospective remedies, and it was left to panels and the AB to resolve this issue.

On the other hand, case law was rather ‘gentle’ when it came to examining whether the requirements for lawful imposition and/or review of duties imposed had been met. A perfectly representative illustration concerns a Swedish claim in US–Swedish Steel Plate. The panel decided that it was not warranted for the US to review 20-year-old duties, even though the following was true:

- Sweden had reduced its production of steel products;
- It was selling more to the EU because it had signed a free trade agreement (FTA) with it (and EU had eliminated duties on imports of Swedish steel);
- Avesta (a Swedish company) had bought a mill in Indiana and, consequently, was selling to the US market through its Indiana site;
- US had concluded a voluntary export restraint with many exporters, as a result of which the health of US industry had considerably improved (§§ 246ff.).

This is a very restrictive interpretation of the “where warranted” standard, which regulated when duties should be reviewed, by any reasonable benchmark.

The combination of the two trends in case law meant that panels would not mess around lightheartedly with impositions, but when they did, they would apply tough standards.

3.2.2 Even Tougher

The AB tightened the screws in several directions in the realm of contingent protection. I have already explained why, in my view at least, the AB set aside the expressed intent of the instigators of the standard of review embedded in Article 17.6(ii) of Antidumping.

The AB made recourse to contingent protection even more onerous through its (correct in my view) understanding of the causality requirement. Anything can affect the trade outcome, and a trigger happy investigating authority is often happy to attribute to dumped- or subsidized- or increased imports the injury inflicted by other factors. Why blame bad investment, for example, when jobs are at stake, and costs can be shifted to foreign producers?

The AB requested evidence of genuine and substantial relationship between cause and effect before recourse to duties has been made. And this time, it meant business. The standard as understood by the AB requires attribution of injury to say increased imports, and evidence of nonattribution of injury to increased imports when other factors have caused it. It requests from WTO members to examine their conclusions in light of other alternative explanations, and decide on the imposition of duties only when they have done so. Sykes (2003a), reading the case law, has gone so far as to ask whether it will ever be possible for an (elaborate) investigating authority to meet the test established in case law? We would add that case law in the realm of nondiscrimination would have been drastically different had the AB adopted a similar understanding of the causality-requirement there as well. Howse (2016) is certainly right in pointing the discrepancy in this direction.¹¹

3.2.3 A Touch of Mild

The AB cannot be accused for consistency. Consistency, of course, is not per se a value, for one can be consistently wrong. Explanations of inconsistencies though, are warranted for courts to retain their credibility. Under the circumstances, it is quite odd that the AB refused to see a causality requirement in sunset reviews. Why is it the case that one needs to show that dumping causes injury in the original investigation, and not so at the sunset-stage? The AB pointed to the absence of specific language to this effect. True. But does the absence of specific language lead to the conclusion that, whereas only injurious dumping can be sanctioned at the stage of original investigation, noninjurious dumping can be counteracted at the sunset-stage? Had the AB investigated even briefly the negotiating record, it would have realized that the introduction of sunset reviews was a hard fought victory for its proponents. The whole idea was that duties lapse, unless a sunset review points to recurrence of injury in case of withdrawal. Negotiators a small window for continued imposition of duties. The AB turned it into a wide avenue.

Indeed, Howse and Staiger (2007) expressed their profound disagreement with this statement, and argued for a comprehensive legal test that could be applied in order to sustain the plausibility of continued impositions. The AB turned a blind eye to this paper, as it consistently does to academic writings that rarely if ever feature in footnotes. As a result, an investigating authority can routinely meet the sunset requirements. While sunset reviews exist only in antidumping and countervailing, the least disturbing (for international trade purposes) form of contingent protection, namely safeguards, is the one that WTO members will have more difficulty in employing.¹²

¹¹ It is true that, on occasion, the AB on the one hand announces a very demanding causality test, only to decide on causality on shaky basis a few paragraphs later. The standard announced though, might still serve as some sort of signaling mechanism to the effect that the AB will take this exercise seriously.

¹² There are other examples of soft behavior by the AB. In the softwood lumber disputes for example, the AB disregarded the explicit wording of Article 14 of SCM, and went on to justify the use of benefit calculation by the US that did not correspond to any of the standards reflected in the exhaustive (on textual grounds, at the very least) list of this provision.

The AB has not pronounced in comprehensive manner on remedies. WTO panels, with one exception, have consistently held that remedies are prospective, that is, the obligation to compensate kicks in from the end of the reasonable period of time during which unsuccessful defendants must bring their measures into compliance with the WTO. This is a major concession to the members that defended this view during the Uruguay round, and did not manage to persuade the rest of the membership on this score.¹³ The AB has implicitly sided with this approach.

So tough yes, the AB has been tougher on trade instruments than it has when adjudicating disputes about domestic policies. But at the end of the day, losers know that they will not be severely punished for breaking the law. Cheap exit for the first five years or so is now de facto institutionalized.

3.2.4 And, as Usual, a Lot of Confusion

There are numerous inconsistencies in case law on trade instruments, and in this respect, the AB reproduces the record of its case law under domestic instruments. To provide but one illustration, I will refer to the case on pass-through, a major issue in subsidies.

The AB dealt in quick succession with two cases regarding pass through, both discussed in Grossman and Mavroidis (2007). The first time it held that payment of market price when goods are privatized always exhausts previously bestowed benefits. In the second, it held that privatization at arm's length could result in exhaustion of similar benefits. Both findings cannot be right, one of them has to be wrong. And yet, the AB did not even bother to address the inconsistency between the two reports. *Errare humanum est*, and it does not diminish the credibility of a court to state that it has erred. The European Court of Justice (ECJ) certainly did not suffer when in its notorious *Keck & Mithouard* judgment took explicitly distance from prior case law.

There is worse. The AB never explained in its second decision under what conditions payment of market price exhausts benefits. And, it did not ask the correct question in neither of the reports, which, as Grossman and Mavroidis (2007) show is whether the investment has (or has not) become inframarginal. The result is predictably, confusion, and there is proof for it. Amazingly, the proof of confusion has been reflected in an AB report. In *EC and Certain Member States-Large Civil Aircraft*, the three members of the division had three different opinions regarding extinction of subsidies as a result of privatization at arm's length when a "fair market value" had been paid (§726).

4. The AB Legacy

Stein (1981) offers this wonderful passage when trying to explain the merits of keeping the Court, a decisive court, away from the public eye:

tucked away in the fairyland Duchy of Luxembourg and blessed, until recently, with the benign neglect by the powers that be and the mass media, the Court of Justice of the European Communities has fashioned a constitutional framework for a federal type structure in Europe.

I have often wondered how the AB members when they go back and revisit their work. What will they think of reports a hundred and twenty pages long only to announce at the end that they could not decide the issue before them? What will they think of the confusion they have created through hundreds of pages where the same issue is discussed from all sorts of unnecessary angles only to confirm a decision they have reached in the first paragraph of the discussion? And what will they think when asking themselves, how much they have contributed into making the WTO a policy-relevant instrument across the membership?

¹³ In Mavroidis (2016a), I discussed this issue at length.

4.1 Political Deference vs. Deference when in Uncertainty

Deference is of course, defensible, when dealing with disputes concerning domestic policies. The typical scenario would be a case where only Home knows the rationale for regulating, and it would have little incentive to inform Foreign about it. If Home reveals the truth (regulation is meant to protect the domestic matador), it goes to jail. If Home lies, then it might be out of jail. It is a classic prisoner's dilemma. The question for the judge will be to devise a test that will provide Home with the incentives to reveal the true information. And even then, uncertainty might still persist. Wise courts would then, in face of uncertainty, avoid committing a false positive, and would defer to the legislator. The fact that the legislation is anyway nondiscriminatory on its face at least, and that the reputation of the institution might suffer if panels go at it hammers and tong, are factors that reinforce the argument for deference.

What did the AB do? None of the above, is the short answer. It devised a "no intent-cum-no trade effects"- standard, that allows it, unconstrained by any methodological discipline, to decide cases brought before it.¹⁴ Under the circumstances, deference that the AB shows is the automatic consequence of the decision by regulators to protect human health. There is nothing more to it.

Let us compare for a moment with the attitude the AB has shown when addressing say antidumping. It is true that the contingent protection instruments reflect an injury standard, and a statutory requirement for demonstration of causality between say dumping and injury. The discipline on domestic instruments does not. But is this reason enough to abandon a serious causality analysis, and replace it by guesswork, and a standard that for all practical purposes amounts to judging on (in-)consistency of challenged measures based on one (out of many) plausible scenarios?

Think of Korea-Various Measures on Beef for a moment. Complainants have not suffered trade damage. The intent, the AB accepted, was Korea's quest for punishing deceptive commercial practices and protecting consumers. What did Korea do wrong? Maybe, the dual retail system could lead to more Korean beef being sold in the future when the quota would have been rescinded. Maybe not. In fact, the latter is more probable, since the Korean law imposed no switching costs if traders of Korean beef were to turn to selling imported beef, which they would probably do when the quota would be rescinded in light of the price differential between imported and domestic beef. Does the probability that they might continue to sell more Korean beef suffice for Korea to be proclaimed the culprit in this dispute?

To avoid misunderstandings, I am not suggesting that the AB should view a measure aiming to protect health in the same way it reviews the legality of antidumping. Assuming the propensity to commit mistakes is the same across WTO covered agreements, errors when a health policy is outlawed can be devastating, whereas in the case of antidumping they will be limited to pure trade damage. I am suggesting though, that, in the name of avoiding costly errors, the AB cannot do away with its self-imposed requirement to issue *reasoned* reports (US-Lamb, §106).

4.2 Keep them Happy

Overall, the members of the AB can pride in that they have not caused a major upheaval. Bar the issue of providing a stage for *amici curiae*, a rather inconsequential decision that surprisingly provoked a heated debate across the membership, no decision by the AB has provoked a (quasi) unanimous hostile reaction. Furthermore, the membership has not voted with its feet. In Mavroidis and Sapir (2015), we provided empirical evidence to the effect that there is no forum diversion. The number of disputes in

¹⁴ Grossman et al. (2013) provide a test that panels could usefully employ to operationalize nondiscrimination.

the second decade of the AB is substantially lower than that in the first decade,¹⁵ but WTO members still litigate only before the WTO. If it ain't broke, no need to fix it then.

Absence of massive opposition to reports (*à la amici curiae*) does not mean that the membership adheres to the findings of the AB. A brief perusal of the DSB minutes where reports are discussed suffices to notice the dissatisfaction of many members with particular decisions.

The point here is that the authority of the AB has not been contested. Disagreements do occur, but no one has requested total recall. The establishment of an exclusive forum to adjudicate disputes was the holy grail of the negotiations though,¹⁶ and rationally trading nations have not been prepared to let go one of their major achievements in the name of poor judgments. One should also not lose sight of the fact that judgments are not costly. Cautiousness when human life/health is at stake avoids disasters. Prospective remedies guarantee at the very least sufficient time to adjust.

In this I side with Ethier (2006) who explained why, in presence of uncertainty about the identity of losers in adjudication, the common incentive is to opt for weak remedies. The AB seems to implicitly at least acknowledge as much.

4.3 Keep them Happy?

Is this however, what the AB should be doing? Should it be happy when it has not provoked a major upheaval? Is this the measure for success?

Horn et al. (2010), and Maggi and Staiger (2011) provide mutually reinforcing accounts why the role of the judge is to 'complete' the contract through case law. Roughly, since trade agreements only through generic language can address the various issues of trade integration, judges by specifying rules to the facts of the case provide information about the coverage of existing disciplines, as well as the manner in which they will be adjudicating future disputes. The judicial function understood in this way is particularly relevant in the WTO context, where renegotiation of agreements is almost an impossibility in light of the consensus-requirement.

To perform this function, judges must privilege methodology over political sensitivity or any other similar concern. Furthermore, the two are not necessarily at odds. As stated above, in face of uncertainty, it is quite rational for a judge to be extra cautious when human life/health is at risk. This should be though the outcome of an intellectual process, a methodology that will attempt to link facts to the relevant legal discipline. This is where the AB has failed, as the examples provided above have attempted to show.

5. Concluding Remarks

The negotiators of the Uruguay round did not spend too much thinking about the AB. It was thought more as a counterweight to the automaticity in establishing panels and adopting their reports, than anything else. Furthermore, it is an entity that operates under severe time constraints, which, as empirical analysis shows,¹⁷ it routinely respects. These are all mitigating factors that one should take into account when discussing the quality of its output.

¹⁵ Note that the WTO counts members in the second decade than in its first.

¹⁶ In Mavroidis (2016a) I discuss the negotiating record in some detail.

¹⁷ Horn et al. (2011). This is quite something, and should not be taken lightheartedly. There is no theory explaining when disputes arise in equilibrium (and, consequently, no theory distinguishing between good- and bad faith disputes), and courts might often be surprised to see cases be submitted to them. On top of this, they must react within a few weeks, and often, they have had no time to think about the issue before them. Case law, assuming it is reliable, can help address in part at least this issue.

One should also take into account the enormity of the task. WTO 'courts' are the busiest state-to-state courts in international relations. The AB was put in place in order to provide guidance regarding the manner in which panels should understand WTO law. It is one case at a time as far as the AB is concerned, and this should mean that it should cross that bridge when it comes to it, and not before. Undeniably though, and by its own admission in its report on US-Stainless Steel (Mexico), panels should cross that bridge in the way it has showed to them, and not in a different way (§§158-162). The AB thus, has the responsibility to ensure that panels will cross bridges in a particular way. This is what methodology amounts to.

The AB seems to neglect all this. The best proof that it has not honoured this task, its main task, is the number of panel reports that it modifies or reverses. It simply cannot be that people do not get it. There are cases of course, where panels have consciously deviated from prior case law established by the AB. A couple of panels voluntarily did not adhere to the outlawing of zeroing by the AB, and explained in plain English why this had been the case.¹⁸ But there must be something wrong when panel reports cite the AB in support of their position, only to be reversed or modified by the AB a few weeks later.

Reasoning in (almost) endogenous manner (what is politically acceptable?), it has failed often enough to provide guidance how to resolve disputes. It is definitely not a coincidence that it routinely overturns specific panel findings even on issues, like nondiscrimination, where case law should be smooth by now in light of the ample opportunity the AB has had to set the record straight. Furthermore, its inability to acknowledge that it has overturned itself perplexes the intended recipients of its output (panels and disputing parties alike) as to which law applies where. Add the hundreds of totally useless pages that hide the few important paragraphs, and you have the whole nine yards.

It is high time for drastic change. Legal drafting should improve. It is simply unnecessary to fill a hundred plus pages only to state that the AB cannot decide the issue. It is equally unnecessary to see the text of the Vienna Convention cited ad nauseam in every report, as it is to read the arguments by parties repeated three and four times. An honest opposition to prior case law dealing with the same issue will greatly help. The Vienna Convention is not surrogate for the missing methodology. It is a checklist of elements that the AB can or may take into account. The methodology that it will employ to adjudicate disputes is a totally different exercise. Judging from the record so far, this will be the Everest that will need to climb in the years to come.

Post Scriptum

As I was finishing this paper, reports emerged to the effect that the US was opposed to the reappointment of the Korean member of the AB.¹⁹ The reason given, according to the reports at least, was that this member of the AB had participated in cases where the body had overstepped its mandate. If true, then this is the first major crisis since the *amici curiae* saga that I have discussed above. It is also a more important crisis by any benchmark, since this will be a first where a WTO member opposes reappointment. The grounds mentioned to justify the decision complete the scenario.

In my view, the US action does not disturb the findings of Howse (2016). The mandate of the AB can be overstepped irrespective of the subject matter of a dispute handled by the AB. I think it will be a pity though, if the US were to limit its dissatisfaction to a person. After all, the individual they want to see leave the body did not sign any dissenting opinions, which anyway are, on paper at least,

¹⁸ Mavroidis (2016), Chapter 2, volume 2 discusses them in detail. There are a few other, more nuanced cases of deviation, like the standard employed in the panel report on EC-Commercial Vessels, when compared to the AB report on US-Offset Act (Byrd Amendment).

¹⁹ Inside US Trade, World Trade on Line, Saturday 14 May, 2016.

anonymous, and it is hard to discern what his influence has been in unanimous decisions that the US obviously does not agree with. Anyway, the AB is by statute collectively responsible for its decisions.

In my view, this action will not solve the problem. It is high time that a serious, institutional debate about the workings of the AB takes place within the WTO. Issues like the corporate governance of the body, the selection process, the background and the role of the Secretariat (the clerks) should not be taboos. So far, the only scrutiny comes from the annual reports written by academics, and is limited to the quality of the output.²⁰ The AB occupies only one provision in the DSU. It deserves substantially more thinking.

²⁰ The American Law Institute (ALI) first put together a group of academics that met annually and discussed the case law of the previous year. This group has continued to meet annually and discuss the case law after the ALI left the project. Papers are annually published in a commercial review, see <http://journals.cambridge.org/action/displaySpecialPage?pageId=7624>

References

- Baldwin, Robert E. 1970. *Non-tariff Distortions in International Trade*, Brookings Institution: Washington, DC.
- Croley Steve, and John H. Jackson. 1996. WTO Dispute Procedures, Standard of Review, and Deference to National Governance, *American Journal of International Law*, 90: 193–213.
- Davey, William J. 2006. Reflections in the Appellate Body Decision in the Hormones Case and the Meaning of the SPS Agreement, pp. 118–32 in George A. Bermann and Petros C Mavroidis (eds.), *Trade and Human Health and Safety*, Cambridge University Press: New York City, NY.
- Ethier, Wilfried J. 2006. Punishments and Dispute Settlement in Trade Agreements, The Equivalent Withdrawal of Concessions, *Keio Economic Studies*, 42: 1–23.
- Grossman, Gene M., Henrik Horn, and Petros C. Mavroidis. 2013. Domestic Instruments, pp. 205-345 in Henrik Horn and Petros C. Mavroidis (eds.), *Legal and Economic Principles of World Trade Law*, ALI, The American Law Institute Reporters' Studies on WTO Law, Cambridge University Press: New York City.
- Grossman, Gene M., and Petros C. Mavroidis. 2007. Recurring Misunderstandings of Non-Recurring Subsidies, pp. 381–90 in Henrik Horn and Petros C. Mavroidis (eds.), *The American Law Institute Reporters' Studies on WTO Case Law*, Cambridge University Press: New York City, NY.
- Horn, Henrik, Johannesson, Louise, and Petros C. Mavroidis. 2011. The WTO Dispute Settlement System: 1995-2010, Some Descriptive Statistics, *Journal of World Trade*, 45: 1107-1138.
- Horn, Henrik, Giovanni Maggi, and Robert W. Staiger. 2010. Trade Agreements as Endogenously Incomplete Contracts, *American Economic Review*, 100: 394–419.
- Howse, Robert. 2016. Global Governance by Judiciary, *European Journal of International Law*
- Howse, Robert, and Robert W. Staiger. 2007. US-Sunset Reviews of AD Duties on Corrosion-Resistant Carbon Steel Flat Products from Japan, pp. 601–621 in Henrik Horn and Petros C. Mavroidis (eds.), *The WTO Case-Law of 2001–3: The American Law Institute Reporters' Studies*, Cambridge University Press: New York City, NY.
- Hudec, Robert E. 1993. *Enforcing International Trade Law*, Butterworths: London, UK.
- Hudec, Robert E. 1998. GATT/WTO Constraints on National Regulation: Requiem for an 'Aims and Effect' Test, *International Lawyer*, 32: 619-647.
- Jackson, John H. 1969. *World Trade and the Law of the GATT*, Bobbs-Merrill: Indianapolis, Ind.
- Irwin, Doug, Petros C. Mavroidis, and Alan O. Sykes. 2008. *The Genesis of the GATT*, Cambridge University Press: Cambridge, Massachusetts.
- Maggi, Giovanni, and Robert W. Staiger. 2011. The Role of Dispute Settlement Procedures in International Trade Agreements, *The Quarterly Journal of Economics*, 126: 475–515.
- Mavroidis, Petros C. 2002. Amicus Curiae Briefs Before the WTO: Much Ado About Nothing, pp. 317–329 in Armin von Bogdandy, Petros C. Mavroidis and Yves Meny (eds.), *European Integration and International Co-ordination*, Studies in Trans-national Economic Law in Honour of Claus-Dieter Ehlermann, Kluwer: Leiden, the Netherlands.
- Mavroidis, Petros C. 2016. *The Regulation of International Trade*, volumes 1, and 2, MIT Press: Cambridge, Massachusetts.
- Mavroidis, Petros C. 2016a. Mind Over Matter, in Kyle Bagwell and Robert W. Staiger (eds.), *Handbook in Commercial Policy*, Oxford University Press: New York City NY.

- Mavroidis, Petros C., and André Sapir. 2015. Dial PTAs for Peace: The Influence of Preferential Trade Agreements on Litigation between Trading Partners, *Journal of World Trade*, 49: 351-374.
- Neven, Damien J., and Joseph H.H Weiler. 2006. Japan—Measures Affecting the Importation of Apples: One Bad Apple? pp. 280–310 in Henrik Horn and Petros C. Mavroidis (eds.), *The WTO Case Law of 2003: The ALI Reporters 'Studies*, Cambridge University Press: New York City, NY.
- Stein, Eric. 1981. Lawyers, Judges and the Making of a Transnational Constitution, *American Journal of International Law*, 75: 1-42.
- Stewart Terence, Susan G. Markel, and Michael T. Kerwin. 1993. Antidumping, pp. 1383–710 in Terence P. Stewart (ed.), *The GATT Uruguay Round: A Negotiating History (1986— 1992)*, Kluwer Law: Deventer, the Netherlands, and Boston, Massachusetts.
- Sykes, Alan O. 2003. The Least Restrictive Means, *University of Chicago Law Review*, 70: 403–416.
- Sykes, Alan O. 2003a. The Safeguards Mess: A Critique of WTO Jurisprudence, *The World Trade Review*, 3: 216–296.
- Wilcox, Clair. 1949. *A Charter for World Trade*, Macmillan: New York City, NY.

Author contacts:

Petros C. Mavroidis

Edwin B. Parker Professor of Law at Columbia Law School

435 West 116 Street

New York, NY 10027

USA

and

Robert Schuman Centre for Advanced Studies, EUI

Global Governance Programme

Via delle Fontanelle 18

I-50014 San Domenico di Fiesole (FI)

Italy

Email: petros.mavroidis@eui.eu