

Comment

Brazil – Measures Affecting Imports of Retreaded Tyres (DS322)

Prepared for the ALI Project on the Case Law of the WTO

J. H. H. WEILER

New York University School of Law

The importance of the Bown and Trachtman analysis of *Tyres* transcends the specifics of the case itself. It is one of the most powerful and explicit indictments of the very competence of the Appellate Body to perform a task at the core of its judicial mission. In terms of its ‘bottom line’, the Appellate Body, it would seem, got it right. The Brazilian scheme did violate its WTO obligations. But, in law, a right result might be a necessary but not a sufficient condition. After all, the Biblical priests administering the test of bitter water to ‘a wife that goeth aside to another instead of her husband’ (Numbers 5: 11–31), or the Delphi Oracles themselves also got it right from time to time. Oracular decisionmaking, the authority of which rests on the status of the decisor, rather than the quality of the reasoning, is antithetical to the judicial function.¹

There are perhaps open questions as to the appropriate legal test to apply in the Articles III, XI, XX cluster. One reason is the problematic drafting of the text itself. One would, to give but one example, expect that some form of a necessity test (however one wants to construe necessity) would be indicated for all exceptions to the core prohibitions to which Article XX applies. And yet the word itself is distributed almost haphazardly in the text of Article XX. Attempts to rationalize the use of the word necessity are ultimately just that – rationalizations. More deeply, there might be a tension between the underlying economic rationale of disciplines such as national treatment in terms of total welfare and the latitude the relevant legal text allows the decisor to pursue that rationale. The analysis of Bown and

¹ In this comment, I sidestep the important if, by this point, tiresome debate about the role of ideology and personal proclivity in a rules-and-principles-based jurisprudence. Here we are addressing a jurisprudence that prides itself on its principled hermeneutics but that Bown and Trachtman largely persuasively show to be ‘incoherent.’

Trachtman justly allows for this even if they are not shy in expressing their own normative preferences.

But it is that very textual impreciseness and tension between policy and text that had the work of the Appellate Body cut out for it: develop over the years a coherent jurisprudence with each new case offering a chance of honing, self-correcting, and at times even changing course (as did, for example, in different circumstances, of course, the ECJ in its shift from *Dassonville* to the *Keck*). There has not been a paucity of cases that would have enabled such a common-law-type jurisprudence to develop. And yet, despite the very centrality of the type of case *Tyres* represents, the Appellate Body is still found to be meandering among the different possible approaches consistent with the language of Article XX. Even more damning, in opting for one or the other, whenever it strays from the false safety that its laughable dictionary hermeneutics supposedly provides, it shows itself, ironically, to be extraordinarily loose with its own use of language (no dictionary will help you give meaning to the formulae pronounced by the Appellate Body itself) and vague (if not naïve) with the economic realities that these formulae entail if they are to be taken seriously. Thus, Bown and Trachtman:

[Whilst the] WTO Appellate Body has spoken of a test that weighs and balances to some degree each of the four factors mentioned above, it has never documented in an opinion its application of this type of test, or insisted that Panels actually apply this type of test. Most importantly, in *Brazil–Tyres*, it has shown itself unwilling to evaluate for itself, or to require a Panel to evaluate, in any but the most gross categories, any of these four factors. Yet, one might ask, if you consider these factors, but you do not evaluate them, in the sense of assessing their magnitude, and you do not compare them with one another – the costs with the benefits – how do you determine which domestic measures are acceptable and which are not? (at 88).

And:

Furthermore, this case is troubling for the failure of the Panel and the Appellate Body to deal effectively with the problems of defining the measure to be evaluated, defining the chosen level of protection, and evaluating the contribution made by the measure at issue. The failure effectively to address these critical components of the analysis makes the assessment of each of the factors addressed in the Appellate Body's analysis imprecise and ultimately malleable.

The result of the combination of these problems is an opinion that is so incoherent as to leave states unsure as to what types of measures may withstand scrutiny. (at 88–90)

Res Ipsa Loquitur.

One sign of the failure to develop a coherent jurisprudence (and another irony for a judicial body that puts so much emphasis on textualism) is the absence of consensual usage of nomenclature. The taxonomic clarification of the authors is welcome but is unlikely to dispel all the lexical fog. The German school will look

in vein for their notion of ‘proportionality’, often found in the literature, and which includes in its three prongs, an examination of legitimacy of purpose (which is taken for granted in an Article XX context), LTRARA, and, finally, as a third factor, cost–benefit balancing. The ECJ and *la doctrine* that shades its jurisprudence, typically use the term ‘proportionality’ to denote only in the narrower LTRARA sense. Even LTRARA and LTIARA are often collapsed into a simple Least Restrictive Measure (LRM) appellation.² They certainly both belong to that family.

Setting aside the semantics, I need to revert to the various tests in order to set up one notable paradox that emerges from the analysis of Bown and Trachtman and that merits some reflection. It is a paradox that is rooted in, and is a manifestation of, an oft-present tension between analytical jurisprudence and legal realism – where the life of law once again shows itself to be based not on logic but experience.

First, I want to draw, in the sharpest possible terms, the philosophical and policy differences between the balancing/proportionality (ECJ version) test on the one hand and LRM tests on the other.

The LRM family derives from, and is meant to be respectful of, a supposedly cardinal, even ‘constitutional’ premise of the GATT³ system: The freedom of Members to set their own level of acceptable risk in their political economy – their so-called ‘regulatory autonomy’. It is a mantra that the Appellate Body repeats again and again, even in *Korea–Beef* itself, here in the context of an Article XX(d) analysis:

It is not open to doubt that Members of the WTO have the right to determine for themselves the level of enforcement of their WTO-consistent laws and regulations. (Recital 176)

This principle would seem to mandate LRM tests and actually prohibit Balancing. LRM respects the level of protection set by a Member. It normally assumes that the measure adopted is an indication of the level of risk the Member wishes to achieve,⁴ and then simply seeks to check whether the measure adopted to achieve that level is the least restrictive in its trade effects. By contrast, to ‘balance’ is to ask an altogether different question – as so lucidly explicated by Bown and Trachtman. It is not simply to ask whether an alternative less-restrictive measure exists, but (even if such an alternative does not exist) whether the benefits, however defined, obtained by the measure adopted by the State outweigh certain costs (however defined) associated with that measure. If the decisor reviewing the case decides that the cost–benefit is disproportionate – typically because the cost

² I find the Least Treaty Inconsistency test (LTIARA) unhelpful analytically – it typically begs the question and requires some LTRARA in order to work.

³ I use GATT advisedly. SPS and TBT may differ from traditional GATT disciplines precisely on this point.

⁴ The issue of ‘consistency’ with other measures, e.g., as referenced in SPS, is an instrument whereby this assumption may be ‘gently’ challenged.

associated with the restriction on trade outweighs the benefits that the measure may produce (e.g., in preventing consumer fraud) and requires for *this* reason that the measure be modified, it is substituting its value judgment for that of the State in determining the level of acceptable risk.⁵ Adopting a balancing test has to mean that, at least in some circumstances, the supposed hallowed right of a State to set its own level of acceptable risk is not absolute.

Korea–Beef is instructive: the Appellate Body says:

It seems to us that a treaty interpreter assessing a measure claimed to be necessary to secure compliance of a WTO-consistent law or regulation may, in appropriate cases, take into account the relative importance of the common interests or values that the law or regulation to be enforced is intended to protect. The more vital or important those common interests or values are, the easier it would be to accept as ‘necessary’ a measure designed as an enforcement instrument. (Recital 162)

But that, of course, is a political value judgment – how to balance risk against benefit – and part of the right about which the Appellate Body seemed to have no doubts. If it is the right of the State to determine the level of enforcement, as stated by the preceding citation from the decision, whence the right of the Treaty interpreter to do this balancing between value and risk for the State?

The Appellate Body adds:

Another aspect [of examining whether an enforcement measure is necessary] is the extent to which the compliance measure produces restrictive effects on international commerce, that is, in respect of a measure inconsistent with Article III:4, restrictive effects *on imported goods*. A measure with a relatively slight impact upon imported products might more easily be considered as ‘necessary’ than a measure with intense or broader restrictive effects. (Recital 163 footnotes omitted)

What seems to emerge clearly from the explicit language of the Appellate Body here and from its actual decisions in other cases, e.g. *Apples*,⁶ which are predicated on a similar balancing approach, is that, for example, if the risk in question is to life and limb, the Appellate Body will countenance and consider necessary very strict enforcement aimed even at zero risk and even if that means a very heavy burden on imports. But if the risk in question is far more benign, e.g. mere consumer fraud, the State will not be allowed measures that produce a similar heavy burden on imports. Only measures with a less restrictive impact on trade will be allowed if the risk of, say, fraud becomes larger.⁷

⁵ Sometimes the balancing may be reviewed as a proxy for a good-faith test. See *infra*.

⁶ *Japan – Measures Affecting the Importation of Apples*, WT/DS245/AB/R 26 November 2003.

⁷ When the Appellate Body, upon rejecting the dual retail system, says, in Recital 178: ‘Therefore, the question remains whether other, conventional and WTO-consistent instruments can not reasonably be expected to be employed to achieve the *same* result’ (emphasis added) it is clear that what is proposed would not achieve the *same* result. Achieving the *same* result by other enforcement mechanisms will entail a far heavier administrative and financial burden; in effect, the burden of the case, which derives from

In other words, there is an inherent tension, even contradiction, between the two passages from Recitals 162/3 and 176 of Korea Beef quoted above.⁸

Let us now couple this clarification with the critique of Bown and Trachtman. It is reminiscent of the famous Jewish joke from the Borscht Belt: the food was bad, and, what's more, there was not enough of it. Simply in moving from the LRM family to Balancing, the Appellate Body would seem to be violating that very cardinal principle, regulatory autonomy, on which it repeatedly insists. And, then, as Bown and Trachtman persuasively demonstrate, it does such a bad job at applying that very test that:

The result ... is an opinion that is so incoherent as to leave states unsure as to what types of measures may withstand scrutiny. (at 89–90)

One would expect that in the face of such an unappetizing meal, the Members would balk. That a jurisprudence which seems to compromise their hallowed right to set their own risk level and, the sloppy and vague application of which, leads to incoherence, would generate at least the same level of political hue and cry that welcomed the Appellate Body's decision on the far less important issue of *Amici* briefs. Or that it would generate appeals, such as the notable appeals by the winning sides(!) in *Japan Alcoholic Beverages* and *Asbestos*, where the gripes with Panel decisions turned not on result but on doctrine and methodology. And yet, *mirabile dictu*, both in the EU in which the slippage from LRM to Balancing was achieved *sub silentio*⁹ and in which its operationalization is really at the same vague intuitive level critiqued by Bown and Trachtman, and in the WTO system itself where these developments have been with us for quite some time, Members have largely reacted with indifference. They simply do not seem to care.

How does one explain this indifference to a flawed jurisprudence in such a central area?

The normative dimension of Bown and Trachtman's critique is interesting. It balks at the vulgar, positivist, legal-realism reflex, which is satisfied in simply stating, with precision, what the law is as articulated by the most authoritative judicial mouthpiece – the Law of England is what the House of Lords says it is. The Constitution of the United States is what the Supreme Court states it to be. In part, their normative critique is based on a quasi essentialists' premise that is rooted in an economic Archimedean point (global welfare), which in their eyes

Recitals 162 and 163, is that preventing fraud at the level achieved by the dual retail system entails, as a matter of balancing, too high a cost on imports when compared to the value in question.

8 Of course, reviewing the mechanisms for enforcing a policy is tantamount to reviewing the actual level of acceptable risk. It is an empty gesture to say that a Member is free to set its own level of acceptable risk but not free to adopt the enforcement mechanisms necessary to ensure that the chosen level of risk is respected. The balancing language used by the Appellate Body in other cases indicates that the vital interest doctrine in *Korea–Beef* is not restricted to Article XX(d) situations.

9 Cf. J. H. H. Weiler, 'The Constitution of the Common Market Place: Text and Context in the Evolution of the Free Movement of Goods', in *The Evolution of EU Law* (Paul Craig and Grainne de Búrca eds., Oxford/New York: Oxford University Press 1999).

should be used as a yard stick for assessing judicial methodology and results.¹⁰ In part, their critique is pragmatic and functional: a jurisprudence that leads to incoherence, fails an even more basic function of the law, namely of giving any guidance on how to act – independently of the question whether the guidance given is right or wrong.

Be that as it may, the Members are the primary clientele of the dispute settlement process. If the jurisprudence of the Appellate Body seems to be acceptable to its primary clientele, can it really be said to be ‘wrong’? The legal-realist claim seems to cut deeper here: the law is so not simply because the Appellate Body has stated it to be so, but because the Members seem, tacitly, to approve of a jurisprudence despite a contradictory doctrinal basis and a deeply flawed methodology. Such a position would surely meet the pragmatic functional dimension of the normative critique and, arguably, maybe even its essentialist dimension. After all, the economic premise of global welfare itself needs an Archimedean point, and can that, in the face of open textured language, really diverge from the view of the authors and ‘masters’ of the Agreement who do not seem to recoil from the doctrine and interpretation that Bown and Trachtman so trenchantly critique? Put differently, if the very test of balancing based on vital interests as articulated in *Korea–Beef* and elsewhere represents such a doctrinal assault on that cardinal freedom – regulatory autonomy to set one’s own level of acceptable risk – and if the application of balancing is so flawed as to lead to incoherence, why are there no meaningful signs of Member reaction and opposition in the political and legal spheres?

Assuming, nonetheless, that there is merit, as I believe there is, in both the doctrinal and methodological critique, I wish to conclude this comment by offering a speculative list of possible reasons for the indifference and equanimity of the Members in the face of such flawed jurisprudence. The reasons offered are not mutually exclusive.

Bown and Trachtman themselves hint at one possible reason: rhetoric of LRM and/or balancing notwithstanding, the reality of Appellate Body practice has been a retreat to the suitability test qualified only by discrimination scrutiny.

A suitability test ... asks simply whether the national measure seems reasonably designed to achieve the purported legitimate goal. It is the least intrusive of these tests with respect to national autonomy. (Ibid.)

If this is the case, it could be understandable why the Members would not balk at the jurisprudence – it seems to grant them, *de facto* at least, maximum liberty. I have doubts as to the persuasiveness of this hypothesis as an explanation for Member reaction (or nonreaction). First, Bown and Trachtman do not, if I read them correctly, claim that the Appellate Body and Panels have in fact retreated to a

¹⁰ I say quasi-essentialist because they will allow that perhaps it is the text of Article XX which constrains the judicial bodies from following their Archimedean point. In principle, they are right, though it is hard to see how the open-textured language of Article XX could be construed as a constraint in this case.

suitability test – a result which could perhaps explain Member acceptance. They are claiming that unless the Appellate Body is willing to address seriously the methodological requirements of, say, magnitude assessment, necessary for a meaningful application of LRM and/or balancing, a suitability test would be the only viable coherent (rather than shambolic) judicial alternative. In fact, *Korea–Beef* itself and other cases (especially cases where the measure in question was rejected) cannot be reasonably explained as an application, even *sub silentio*, of the suitability test. They are poor applications of balancing – of which *Tyres* itself is a prime exhibit. Moreover, it is not even clear (again as Bown and Trachtman intuit) that Members themselves would want the suitability-test option, express or implied – at least to the extent that it represents a retreat from LRM. A scrutiny of the cases will show that LRM contestation is an ubiquitous (and perhaps even favorite) tool in the claimant arsenal precisely because of the balance it achieves between deference to another Member’s regulatory autonomy coupled with a rejection of measures that – on the very terms set by the importing State – *unnecessarily* exclude imported products. Members do not only import and regulate, they also export and face regulation of others. Suitability alone might appear to deny too many of the benefits that a robust national-treatment regime would seem to promise.

Another possible reason for Member State indifference to the flaws represented by *Tyres* may be rooted in the exceptionality of WTO dispute settlement in the overall scheme of State regulation and a disconnect between regulatory agencies within the State and the trade agencies. A robust version of this hypothesis would run as follows. WTO dispute settlement provides very little incentives to Members to disseminate among, and internalize within, the myriad State regulatory agencies the strictures of national treatment ex, say, Articles III and XX. Unlike the EC regime where a violation by the State of, say, the equivalent of Article III(2) may result in an obligation of multimillion retroactive tax payback and in which there is even a (somewhat creaky) regime for damages for violating obligations, a WTO finding of violation at most results, at least as tacitly agreed among the Members today, in a prospective obligation to remove the violation. Corrective measures themselves must be the subject of *de novo* dispute settlement, and the entire process before one faces sanctions (which often are less than compelling) may last for years. Although in the small pond of WTO dispute settlement, one is impressed by the growing number of cases, these numbers are infinitesimally small in relation to the vast regulatory matrix of the modern state. In all but the most exceptional, public high-visibility cases (such as *Hormones*, *GMOs*, *Gambling*) a ruling against a measure, even if based on flawed methodology, is but a fleabite. Compromising the regulatory autonomy of the State, partially removing its hallowed freedom to set its own level of acceptable risk, may loom as a hugely sensitive issue in the classroom or at the Centre William Rappard. But, in fact, the real-life specific gravity of these decisions is slight. The Members will continue, unperturbed, to set their level of acceptable risk in their regulatory regimes – and from time to time

face a challenge. The damning statement by Bown and Trachtman, that *Tyres* represents a jurisprudence ‘that is so incoherent as to leave states unsure as to what types of measures may withstand scrutiny’ is simply of no consequence because, in the vast reality of State regulation, the question of whether a measure will or will not withstand scrutiny is typically not on the radar, and even when it is, is of far less consequence than our grave analyses would suggest. The same type of reasoning would apply, *mutatis mutandis*, to the opposite case, where the Panel and the Appellate Body, employing flawed methodology, fail to find a violation.

A third possible reason, and the one I favor most, rests on two planks. Though States wish to maintain the principle of freedom to set their acceptable level of risk, in the vast number of cases, the regulatory agencies apply a similar vague, approximate, and crude assessment of the welfare-implication tradeoff of regulations that restrict trade, as does the Appellate Body, and that is the subject of the critique of Bown and Trachtman. Moreover, very often a measure is adopted without much concern for, or reflection about, the level of risk that it implies. In reality, the level of risk is deduced from the measure rather than measure adopted as a function of a deliberative process of setting an acceptable risk level. State public authorities will, thus, know from their own practice that measures represent a risk level that may be haphazard, often excessive and indefensible. The doctrinal *status quo* is satisfactory because it retains *de jure* regulatory autonomy, but *de facto* allows balancing scrutiny to root out indefensible, haphazardly set, risk levels. Does it really make sense to stipulate zero or a very low level of tolerance to any consumer confusion if the cost is excluding totally or seriously impeding an imported product from the market place (*Korea–Beef*; *Cassis de Dijon*)? On this reading, the only real vice of the Appellate Body was to be more transparent in what it does than its European counterpart, which does the same thing without ‘fessing up’. The methodological *status quo* (flawed as Bown and Trachtman show it to be) is acceptable, because it is a reflection of similar approximations in internal decisionmaking. After all, it is not as if we are seeing the kind of empirical analysis predicated by Bown and Trachtman as a regular feature in the submissions of parties to Panels and the Appellate Body. Complaints are mostly filed and argued, and defenses are mounted, with an analytical apparatus that is oft not all that more sophisticated than the decisions of the Panels and Appellate Body. The problem rests partly with the technical difficulties in adjudicating cases, indeed preparing them, in the methodologically respectable manner articulated by Bown and Trachtman. But it rests also in the overall regulatory culture that accepts such, and of which the WTO decisors are but a mirror.

Even if this might be the case, or at least part of the explanation, this does not ‘excuse the sloppiness of the Appellate Body’ as articulated by the biting critique of Bown and Trachtman. They are right to take the Appellate Body to task: part of the task of the WTO disciplines (expressly so in SPS, but implicitly elsewhere) is to root out this type of sloppy policy setting because of its negative impact on welfare.