

Do you mind if I don't smoke? Products, purpose and indeterminacy in *US – Measures Affecting the Production and Sale of Clove Cigarettes*

TOMER BROUDE*

Hebrew University of Jerusalem

PHILIP I. LEVY**

The Chicago Council on Global Affairs

Abstract: The AB Report in *US–Clove Cigarettes* confirmed the Panel's conclusions that imported clove cigarettes were 'like' US menthol cigarettes, and accorded 'less favourable treatment' for the purpose of national treatment under Article 2.1 TBT, but it employed significantly different logic. The AB's reasoning extends earlier GATT jurisprudence into the TBT, applying competition-oriented analysis to the question of product definition, while reserving consideration of regulatory purpose to the comparison of treatment. We consider this emphasis from the perspectives of legal and economic indeterminacy, which we find run in parallel to each other. In the particular questions of the appeal, the AB's decisions are sensible from both perspectives, but ultimately do not add significantly to determinacy. We find it impossible to ignore regulatory purpose in discrimination cases, and we provide a novel economic model for analysing regulatory utility in this respect. In our view, there is not much economic difference in analysing regulatory purpose separately from more observable market considerations, but this sequencing does add political logic, analytical focus and formal transparency, all of which may enhance the legitimacy of WTO dispute settlement rulings, although ideally, further specification from the WTO Membership regarding the methodological content of the national treatment discipline would be preferred.

1. Introduction: the continuing story of 'like products' and regulatory purpose

For more than a generation¹ the GATT/WTO system has been struggling with the problem of distinguishing between protectionist discrimination and legitimate

* Email: tomerbroude@gmail.com.

** Email: plevy@thechicagocouncil.org.

The authors are grateful for comments provided by Barry Appleton, Lorand Bartels, Claus-Dieter Ehlermann, James Flett, Rob Howse, Simon Lester, Petros Mavroidis, Damien Neven and Jasper Wauters. Broude thanks David Amos for diligent research assistance. All views and errors are our own.

¹ We note, as indicative in this respect, albeit with respect to internal taxation, the GATT disputes in GATT Panel Report, *United States–Measures Affecting Alcoholic and Malt Beverages*, DS23/R, 19 June

regulation. Once confined to the interpretation and application of the terms ‘like products’ and ‘less favourable treatment’ in Article III GATT, all the familiar dilemmas have now carried over, *en banc* as it were, but not unexpectedly, to the same terms in the WTO Agreement on Technical Barriers to Trade (TBT). *US–Clove Cigarettes* is one of the trio of major disputes relating to the interpretation and application of the TBT in which separate Panels issued reports during 2011, including some mutually inconsistent legal analyses of the same provisions in the agreement² that had previously escaped adjudication. All three disputes were appealed, affording the Appellate Body (AB) an exceptional opportunity to create consistency in uncharted territory (relatively and judicially speaking), through its 2012 case-law. Arguably, the AB has succeeded in doing so (Zhou, 2012; Marceau, 2013), although some of its findings, including those discussed below, have attracted significant criticism (Mavroidis, 2013).

In this context, *US–Clove Cigarettes* stands out among the three TBT appeals for several reasons. Chronologically, it was the first decision to be released by the AB, and its reasoning therefore coloured and served as grounding for that of the other two AB reports. In substance, it was the only one of the three disputes in which the meaning of ‘like products’ in the non-discrimination clause of Article 2.1 TBT, so central to the identification of protectionist measures, was one of the issues under appeal. In contrast, in *US–Tuna II (Mexico)*, the Panel addressed the question of product likeness merely *sua sponte*, and its findings were not even appealed. In *US–COOL*, the question of like products was not seriously contested by the Respondent at the Panel stage, and was also not subsequently appealed.

Furthermore, *US–Clove Cigarettes* was the only one of the three TBT disputes in which the analysis of ‘less favourable’ treatment, the second leg of Article 2.1 TBT, included an explicit discussion of the product scope for the requisite comparison between domestic and imported products. It was also the only one of the three disputes in which the challenged measure was a product ban as opposed to a labelling requirement. Finally, in contrast to the other two disputes, the Panel in *US–Clove Cigarettes* did not find a violation of the requirements of Article 2.2 TBT

1992, BISD 39S/206, and GATT Panel Report, *United States–Taxes on Automobiles*, DS31/R, 11 October 1994, unadopted, and the GATT Working Party Report, *Border Tax Adjustments*, BISD 18S/97, adopted 2 December 1970. Report of 1970 (‘BTA Report’), which is the benchmark for three of the four ‘traditional’ tests of product ‘likeness’: end-uses, consumer tastes and habits, and product qualities, nature and quality (now commonly referred to as ‘physical characteristics’). The BTA Report noted that ‘the term “like or similar products” caused some uncertainty and that it would be desirable to improve on it’.

² Appellate Body Report, *United States–Certain Country of Origin Labelling (COOL) Requirements*, WT/DDS384/AB/R, WT/DS386/AB/R, 29 June 2012; Appellate Body Report, *United States–Measures Concerning the Importation, Marketing and Sale of Tuna and Tuna Products*, WT/DS381/R, 16 May 2012; and Appellate Body Report, *United States–Clove Cigarettes Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/AB/R, 4 April 2012. See Howse and Levy (2012) for a comparative discussion of the three Panel reports.

regarding the trade-restrictiveness of a TBT measure, and indeed the AB was not requested to address this issue at all.

Thus, the AB report in *US–Clove Cigarettes* provides us with a relatively schematic, uniquely compartmentalized, and non-hypothetical opportunity to focus on the legal and economic difficulties involved with product definition and regulatory purpose involved in findings of *de facto* discrimination in international trade law in general and the TBT agreement in particular. To be sure, product definition is perhaps the most nebulous and persistent of problems in GATT/WTO law,³ characterized by a quest for a holy grail of systematic distinctions between national, facially origin-neutral, measures that are in essence protectionist, and measures that are otherwise 'regulatory' or 'legitimate'. Much of this debate has focused on efforts to distinguish between purportedly objective 'economic' or competition-oriented tests, on the one hand, and more subjective tests that look at the sincerity of the measure's regulatory purposes – the classical problem of the 'aims and effects' test⁴ – on the other hand. *US–Cloves Cigarettes* falls squarely into this problem.

Inevitably, much of the general structure of our analysis is similar to prior critiques of product definition and regulatory purpose in the GATT.⁵ Our contribution to this ongoing debate, focusing upon the *US–Clove Cigarettes* AB Report but intended to go beyond its four corners, embraces what we perceive as the inherent indeterminacy of product definition, an indeterminacy which we find is shared by both legal and economic analysis. We believe this parallel indeterminacy should inform the way we think about a number of specific issues that arose in *US–Clove Cigarettes*, and that are current concerns in the WTO, such as the relative role of economic analysis in dispute settlement and rule-making in the WTO, the application of 'traditional' yet economically grounded tests of product likeness such as end-uses and consumer tastes and habits, the relationship between national treatment in the TBT and in the GATT, and perhaps most importantly, the place of regulatory purpose in the legal analysis of allegedly discriminatory measures.

In Section 2 below we provide a brief, stylized summary of the relevant facts and findings of the Panel and AB in the dispute. In Section 3 we generally discuss and compare problems of indeterminacy in law and in economics in the context of the competition-oriented approach embraced by the AB with respect to product definition. In Sections 4 and 5 we explore two elements of the 'traditional' likeness tests that the *US–Clove Cigarettes* appeal focused on – end-uses and consumer tastes and habits – to demonstrate this correlation of indeterminacy in law and

³ And well before the GATT; the term 'like product' has been traced back to 1794; see Choi (2003: 256). Indeed, Choi's study shares much of the ground covered here, albeit focused on the GATT.

⁴ Flett (2013); Hudec (1998).

⁵ Such as Regan (2006).

economics. In Section 4 we argue that the AB's confirmation whereby end-uses of products relate to their *capability* to be used in a certain way rather than their use in practice is consonant with a broad economic approach, but ultimately does not provide greater clarity or legal determinacy, nor can it be free from regulatory policy considerations. In Section 5, focusing on regulatory purpose and consumer preferences, we seize the bull by the horns, and ask whether legitimate regulatory distinctions should be taken into account in the determination of 'like products' (as per the *US–Clove Cigarettes* Panel, as will be explained below) or rather as part of the analysis of 'less favorable treatment' (as per the AB), arguing that from both legal-interpretative and economic perspectives there is ultimately no significant difference between the approaches. A strictly economic approach would just as well internalize regulatory purpose, as the Panel did. However, the AB's approach is analytically attractive and more transparent, by separating regulatory purpose from more objectively observable considerations. In the concluding Section 6 we consider how our discussion of the indeterminacy of product definition and its relation to regulatory purpose reflects upon the relationship between national treatment under Article III:4 GATT, on one hand, and under Article 2.1 TBT, on the other hand; and on the development of more specific rules on non-discrimination in WTO dispute settlement and negotiations.

2. Relevant facts and findings of the AB in *US–Clove Cigarettes*

2.1 *The measure*

The measure at issue in *US–Clove Cigarettes* was Section 907(a)(1)(A) of the United States Federal Food, Drug and Cosmetic Act (FFDCA) as amended by the Family Smoking Prevention and Tobacco Control Act (FSPTCA), which banned the inclusion of a wide range of artificial or natural flavours (such as cloves, strawberry or vanilla), as constituents or additives, in cigarettes in the US. The measure explicitly excluded tobacco and, more importantly, menthol flavouring, from the ban. The stated objectives of the FSPTCA refer to child smoking as a paediatric disease that creates tobacco dependency among adolescents and millions of future adult smokers. It was also noted that restrictions on marketing and advertising had not been sufficiently successful in curbing the phenomenon, and so a more 'comprehensive' approach – presumably additional regulation of access to tobacco products – was required.

Indonesia produces clove cigarettes and prior to the ban exported them to the US; it does not produce a significant amount of menthol cigarettes. In contrast, the US has a major menthol cigarette industry, but no significant clove cigarette production. It also imports menthol cigarettes. In these circumstances, Section 907 (a)(1)(A) FFDCA posed all the classical questions of discrimination in international trade. It was not discriminatory *de jure*. It was driven by a policy justification – public health – that was not related to international trade. Yet it operated

differentially with respect to certain domestic products and certain foreign products. Should the ban be allowed? And does contemporary WTO law in general, and the TBT agreement in particular, provide us with sufficiently clear rules to decide?

2.2 *The complaint and Panel findings*

At the WTO, Indonesia claimed that the measure was a *de facto* violation of Articles 2.1 TBT (national treatment), 2.2 TBT (more trade restrictive than necessary) as well as Article III:4 GATT (national treatment), arguing that imported clove cigarettes were 'like' domestic menthol cigarettes, and that the selective ban constituted less favourable treatment. The US countered that clove and other flavoured cigarettes were not at all 'like' menthol cigarettes primarily on a policy basis, by arguing that clove cigarettes were attractive to youths experimenting with cigarette smoking, whereas menthol cigarettes were used by all age groups, the majority of whom were already nicotine addicts. Consequently, it argued, unavailability of menthol cigarettes due to their inclusion in the ban would cause mass withdrawal symptoms, which would place an immeasurable burden on the US health care system and create an illicit market for menthol cigarettes.

The *US–Clove Cigarettes* Panel found that there had been a violation of Article 2.1 TBT; that Indonesia had not met the burden under Article 2.2 TBT; and exercised judicial economy with respect to the GATT claim. Moreover, in applying Article 2.1 TBT in its determination of product likeness, it used the traditional GATT criteria of physical characteristics, end-uses, consumer tastes and habits, and tariff classification, but subject to a judicially novel interpretation that distinguished national treatment under GATT from national treatment under TBT. According to the Panel (para. 7.244, *US–Clove Cigarettes* (Panel)), the evaluation of likeness in the TBT agreement must 'bear in mind that the measure at issue is a technical regulation with the immediate purpose of regulating cigarettes having a characteristic flavor, with a view to attaining the legitimate objective of reducing youth smoking'. In other words, according to the Panel, in the TBT, the regulatory purpose of the measure must be directly factored into the analysis of product likeness. Regardless, the end-use of both clove and menthol cigarettes was found by the Panel to be simply 'to be smoked'. More significantly, as far as direct consideration of regulatory purpose is concerned, the Panel found that the relevant group of consumers whose tastes and habits should be evaluated was the group of consumers whom the measure addresses in the pursuit of its stated purpose, namely youth and potential youth smokers. In practice, the Panel determined that both clove and menthol cigarettes appeal to youth for the purpose of starting to smoke (para. 7.232, *US–Clove Cigarettes* (Panel)) and that the products are therefore 'like' in that respect too.

The Panel subsequently found that the selective ban constituted 'less favourable treatment' in that it altered the competitive relationship between domestic menthol cigarettes and imported clove cigarettes to the detriment of the latter, and that the

US explanation for the distinction, namely, the potential impact on US health care systems and the possible development of a black market, were related ‘in one way or another to the costs that might be incurred by the United States were it to ban menthol cigarettes’, while at the same time imposing costs on producers in Indonesia, and therefore could not constitute a legitimate regulatory distinction (para. 7.289, *US–Clove Cigarettes* (Panel)), because they were not unrelated to the origin of clove cigarettes.

The Panel also found that the US had violated Article 2.12 TBT by allowing an interval of less than six months between the publication of the measure and its entry into effect, having found that para. 5.2 of the Doha Ministerial Declaration is a subsequent agreement between WTO Members within the meaning of Article 31(3) (a) of the Vienna Convention on the Law of Treaties, defining the meaning of the term ‘reasonable interval’ in Article 2.12 TBT. This is an interesting finding relating to legal interpretation, subsequently upheld, for different reasons, by the AB, but we will not address it here as our focus is on product definition and regulatory purpose. We turn now to the findings of the AB in this regard.

2.3 *The Appellate Body findings*

Excluding the aforementioned Article 2.12 TBT issue, in the appeal lodged by the US, the AB had to address two issues, which together compose the substance of the national treatment requirements of Article 2.1 TBT:

- (a) *Like products* – Was the Panel correct in its finding that clove and menthol cigarettes were ‘like products’ within the meaning of the TBT, with particular reference to the analysis of ‘end-uses’ and ‘consumer tastes and habits’?
- (b) *L Less favourable treatment* – Should the Panel have included other products in the comparison it conducted, as well as market patterns at times other than the measure’s entry into effect, and should it have accepted the reasons submitted by the US for the distinction between clove and menthol cigarettes as legitimate?

Some of the US claims upon appeal were framed as inconsistencies with the ‘objective assessment’ of facts requirement under Article 11 DSU, e.g., that the Panel should have included in its analysis evidence relating to consumers who were not youth or potential youth smokers, and should not have ignored potential costs of the measure to US entities. In passing, we note that this reflects an increasing use of Article 11 DSU in appeals that in effect aim to challenge factual findings by Panels, which in principle cannot be challenged at the AB level, a trend which merits separate attention. However, we will treat these claims on their merits, inasmuch as they concern the question of determinacy in product definition and regulatory purpose.

It is important to note that in ‘bottom line’ terms, the AB upheld the findings of the Panel (and dismissed all the claims under Article 11 DSU). Thus, the AB too found that the US had acted inconsistently with Article 2.1 TBT, confirming that clove and menthol cigarettes were ‘like products’ and that the former had been

accorded less favourable treatment by the measure, within the meaning of that provision. However, the AB was critical of the Panel's reasoning, and used its report as a platform for setting out, in essence for the first time, its perspective on the interpretation of Article 2.1 TBT (primarily in paras. 84–102 of the AB Report).

The differences between the analytical and interpretative frameworks employed by the Panel, on one hand, and the AB, on the other hand, boil down to two elements.

First, where the Panel was of the opinion that national treatment under the TBT and under the GATT were different, stating explicitly that 'it is far from clear that it is always appropriate to transpose automatically' the GATT approach to the TBT (para. 7.99, *US–Clove Cigarettes* (Panel)), the AB adopted a deliberately harmonious approach that whittles down potential gaps between the non-discrimination rules in the two agreements.

Second, and as a consequence, the AB disagreed with the Panel's distancing from a competition-oriented approach to product definition and the inclusion of the regulatory purpose of the technical regulation as an overarching consideration in the assessment of product 'likeness'. The AB chose to defer the direct consideration of regulatory distinctions to the comparative evaluation of 'less favourable treatment'. Despite these significant differences in legal construction, the AB did not overrule the Panel's findings, generally on the basis that the application of its own legal frameworks to the facts determined by the Panel would not have generated different results. We now set out in some more detail the relevant interpretations and findings of the AB.

2.3.1 *General interpretative construction of Article 2.1 TBT*

Article 2.1 TBT provides that:

Members shall ensure that in respect of technical regulations, products imported from the territory of any Member shall be accorded treatment no less favourable than that accorded to like products of national origin.⁶

This is of course very similar to the text of the first sentence of Article III:4 GATT:

The products of the territory of any contracting party imported into the territory of any other contracting party shall be accorded treatment no less favourable than that accorded to like products of national origin in respect of all laws, regulations and requirements affecting their internal sale, offering for sale, purchase, transportation, distribution or use.

Nevertheless, the interpretative context of the two provisions is different. For example, the Panel attached great importance to the fact that Article 2.1 TBT

⁶ For simplicity's sake we have deleted the final words of the provision, which engage with most-favoured nation (MFN) treatment, that was not directly at issue in *US–Clove Cigarettes* (see AB Report, para. 87).

applies only to technical regulations as defined in the TBT agreement, while Article III:4 GATT applies to a much broader range of measures (para. 7.106, *US–Clove Cigarettes* (Panel Report)). In this respect, the Panel cited back to the AB’s vivid image of an ‘accordion’ in its 1996 *Japan–Alcoholic Beverages* Report,⁷ in order to support the finding that product definitions for national treatment under the TBT and under the GATT could and indeed should be different.

The Panel also accorded weight to the recognition of Members’ right to pursue legitimate, non-protectionist policy objectives in their technical regulations, as reflected in the 6th preambular recital and Article 2.2 TBT, as the very basis for a policy-based approach to ‘likeness’ or product definition, as opposed to a competition-oriented approach, thus distinguishing national treatment under the TBT from established jurisprudential approaches to national treatment under Article III GATT.

In contrast, the AB essentially overruled the Panel’s approach, emphasizing the TBT’s 2nd preambular recital (*Desiring* to further the objectives of GATT 1994) as establishing that ‘that the two agreements overlap in scope and have similar objectives’ (para. 91, *US–Clove Cigarettes* (AB)). To the AB, the right to regulate reiterated in the 6th preambular recital and Article 2.2 TBT must be read together with the 5th preambular recital as well as Articles 2.1–2.2 TBT, that reflect the goal of trade liberalization. Taken together, the TBT agreement balances between the desire to avoid unnecessary obstacles to trade against the right to regulate in a way that ‘is not, in principle, different from the balance set out in the GATT 1994, where obligations such as national treatment in Article III are qualified by the general exceptions provision of Article XX’ (para. 96, *US–Clove Cigarettes* (AB)). This very different interpretative approach erodes the potential gaps between the TBT and the GATT, and served to inform the AB’s analyses of product ‘likeness’ and ‘less favourable treatment’.

2.3.2 *Like products*

Thus, the AB rejected the Panel’s recourse to a policy-based ‘likeness’ analysis, maintaining instead the rough competition-oriented approach traditionally applied to national treatment under Article III GATT. It did so, making light of the Panel’s additional contextual toehold: that the TBT lacks the general statement of competitive relationship regarding national treatment, that the GATT finds in Article III:1 and the note *ad Article III:2* (relating to fiscal measures), with reference to ‘directly competitive or substitutable products’. To the AB, ‘Likeness’ in Article 2.1 TBT ‘is based on the competitive relationship between and among the products’, not on the policy objectives of the measure, even though it may take into account ‘the regulatory concerns underlying a technical regulation, to the extent

⁷ *Japan–Taxes on Alcoholic Beverages*, WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R, adopted 1 November 1996, DSR 1996: I, 97, at 21; see the quotation from the AB at note 9.

that they are relevant to the examination of certain likeness criteria and are reflected in the products' competitive relationship' (para. 156, *US–Clove Cigarettes* (AB)). This analysis is in keeping with the AB's approach in *EC–Asbestos* and subsequent cases relating to Article III:4 GATT (e.g., the fact that one product is carcinogenic and another is not may be the basis for a regulatory distinction, but it is also a physical characteristic and/or an influence on consumer preferences from a competition perspective)(see also paras. 116–119, *US–Clove Cigarettes* (AB)).

With this competition-oriented, GATT-friendly interpretative approach, the AB then reviewed the particular US claims on appeal against the Panel's findings of 'likeness'. With respect to end-uses, it agreed with the US that relying on 'to be smoked' as a common end-use of clove and menthol cigarettes was too general, not recognizing the multiplicity of actual and potential end-uses of products (in this case, the satisfaction of nicotine addiction, the creation of a pleasurable experience from the taste and aroma, and social experimentation). However, the AB emphasized the potential rather than actual end-uses of products: 'what matters in determining a product's end-use is that a product is *capable* of performing it, not that such end-use represents the principal or most common end-use of that product' (para. 131, *US–Clove Cigarettes* (AB)). On this basis, and in reference to some of the Panel's factual findings, the AB found that the more specific end-uses still support a finding of likeness.

Similarly, regarding consumer tastes and habits, the AB disagreed with the Panel's decision to focus only on preferences of youth and potential youth smokers, as derived from the measure's stated regulatory purpose, and found that the Panel should have addressed a larger group of consumers. However, in what might appear to be an acrobatic move, the AB then stated that product substitutability need not be found in all market segments. On the facts of the dispute, it then determined that 'the degree of competition and substitutability . . . for young and potential young smokers is sufficiently high to support a finding of likeness' (para. 145, *US–Clove Cigarettes* (AB)). Thus, the AB found that clove and menthol cigarettes were 'like' for Article 2.1 TBT purposes, but based on a difference of interpretative framework than that of the Panel. We return to a critical analysis of these findings in Sections 4 and 5 below, respectively.

2.3.3 *Less favourable treatment*

The appeal in *US–Clove Cigarettes* also gave the AB an opportunity to set out its general interpretative approach to the 'no less favourable treatment' standard in Article 2.1 TBT. It did so without substantial disagreement with the Panel. In keeping with prior interpretations of the same phrase in Article III GATT (and in particular para. 100, *EC–Asbestos* (AB)),⁸ less favourable treatment exists when a

⁸ Appellate Body Report, *European Communities–Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, 5 April 2001.

measure causes detrimental impact to the foreign like product's competitive opportunities, subject to the possibility that legitimate regulatory distinctions may be drawn between products that are otherwise like. The "treatment no less favourable" requirement of Article 2.1 [prohibits] both *de jure* and *de facto* discrimination against imported products, while at the same time permitting detrimental impact on competitive opportunities for imports that stem exclusively from legitimate regulatory distinction' (para. 175, *US–Clove Cigarettes* (AB)). Additionally, '[I]n making this determination, a Panel must carefully scrutinize the particular circumstances of the case, that is, the design, architecture, revealing structure, operation, and application of the technical regulation at issue, and, in particular, whether that technical regulation is even-handed, in order to determine whether it discriminates against the group of imported products' (para. 182).⁹

In practice, this interpretation had little effect on the issues under appeal. The US first argued that the Panel conducted a comparison of treatment across too narrow a product scope: only Indonesian clove cigarettes and US menthol cigarettes, when it should have compared the treatment accorded to the broader group of domestic products (including other flavoured cigarettes that were banned) and imported like products (including foreign menthol cigarettes, which were not banned). In this respect, the AB upheld the Panel's finding, emphasizing that the comparison should in principle be conducted between the imported products from the complainant and domestic products as a group – but there had been no evidence of sizeable production in the US of non-menthol flavoured cigarettes at any time. Furthermore, the US argued that the comparison should have related to the market as it existed at a time earlier than the ban's entry into effect (which would have included some other US non-menthol, flavoured cigarettes, that had subsequently been taken off the market). The AB rejected this argument as moot – the temporal statements of the Panel related not to the comparison between products, but to the findings of detrimental effect.

Most importantly, in this respect, the US argued that the detrimental effect on competitive opportunities for imported clove cigarettes could be justified by legitimate regulatory distinctions and factors unrelated to their foreign origin, essentially retrying the arguments that had failed before the Panel regarding the potential impact of a menthol ban on US health care systems due to withdrawal treatment,

⁹ The AB, in an extended footnote (n. 372 at para. 179, *US–Clove Cigarettes* (AB)), explicitly rejected an interpretation offered by the US referring to *Dominican Republic–Measures Affecting the Importation and Internal Sale of Cigarettes*, WT/DS302/AB/R, adopted 19 May 2005 (*DR–Cigarettes* (AB)), para. 96, according to which in finding less favourable treatment under Article III:4 GATT, a Panel must additionally verify that the detrimental effect caused by a measure is unrelated to the foreign origin of the product. In our view, this does not reflect on the question of the role of regulatory purpose in national treatment under the GATT (and certainly not under the TBT). The AB merely clarified that the situation in *DR–Cigarettes* related to causation of the detrimental effect and its attribution to the measure, not to its rationale.

and the potential creation of an illicit market. Here the AB ultimately agreed with the Panel's conclusion, on a somewhat more detailed analysis, first examining the market data on the record that suggested discrimination against the group of like Indonesian products, then concluding (quite speculatively) that the withdrawal/illicit market arguments do not stick, because 'it is not clear' that the continued existence of regular, non-flavoured, tobacco products would not prevent the materialization of the risks claimed by the US (para. 225, *US–Clove Cigarettes* (AB)).

2.4 *Summary of findings*

In sum, in terms of legal outcomes, the AB confirmed the Panel's findings whereby Section 907(a)(1)(A) FFDCA is inconsistent with Article 2.1 TBT because it accords less favourable treatment to Indonesian clove cigarettes than it does to US menthol cigarettes, which were deemed to be 'like products'.

In terms of legal reasoning, the AB's approach to national treatment in Article 2.1 TBT is to adopt, in essence, the jurisprudence regarding Article III:4 GATT and to apply it to technical regulations. From an interpretative viewpoint, this approach seems well grounded, and has the advantage of creating coherence between these two central agreements on trade in goods. However, some have argued that this empties the TBT of its special attributes (Mavroidis, 2013). We are not convinced that this is the case, because the TBT contains many additional substantive and institutional provisions that go beyond national treatment. Moreover, the TBT ensures that technical regulations as it defines them are caught in the net of non-discrimination, where they might not have been caught under GATT. This is not the case in *US–Clove Cigarettes*, because the measure could clearly have come within the ambit of Article III:4 GATT, but certain other measures, e.g., labelling requirements or so-called 'non-product related PPMs', would not be dealt with under Article III:4 GATT without significant controversy, at least not at the time the TBT agreement was drafted in the 1990s.

Furthermore, the AB's construction of national treatment in Article 2.1 TBT strikes a different analytical sequence and balance between the goals of economic liberalization and regulatory autonomy. Where the Panel saw (at least in theory) the regulatory purpose of the measure at issue directly informing the identification of 'like products', to some extent diluting the competition-oriented aspect of the term, the composite test employed by the AB entails an emphasis on *competitive* relationship between products in the 'like product' analysis (and also in 'less favourable treatment'), while deferring direct recourse to *regulatory* distinctions to the second stage of 'less favourable treatment' analysis.

We now turn to our main theoretical argument, which will inform our analysis of particular issues in *US–Clove Cigarettes* (AB): that the questions of product 'likeness' and 'less favourable treatment' are characterized not only by legal indeterminacy, but also by economic indeterminacy.

3. Indeterminacy in legal and economic analysis of product definition

3.1 *The problem of legal indeterminacy in national treatment*

3.1.1 *The systemic resignation to the indeterminacy of national treatment*

The differences between the Panel and AB in *US–Clove Cigarettes* in the interpretative construction of national treatment under Article 2.1 TBT underscore the relative legal indeterminacy of the normative discipline of national treatment. With all the contextual make-up, Article 2.1 TBT, like Article III:4 GATT, provides us with only two, barebones, terms to work with: ‘like products’ and ‘less favourable treatment’. To begin with, this language provides little guidance regarding application. The customary international law rules of interpretation mandate recourse to the ordinary meaning of the text in good faith and in the light of their context, object and purpose – all of which are rather open to debate, as the AB’s disagreement with the Panel’s approach makes evident. The prior development of national treatment jurisprudence under Article GATT III actually complicates things further: should Article 2.1 TBT be interpreted identically, or should some weight be attached to the different context? Clearly, the AB disagreed with the Panel’s choice on this issue, but notably, was only moderately critical of the Panel in this (and in every other) respect. The AB in this dispute never trashes the Panel’s approach, respecting it as reasonable, yet disagreeing with it. And indeed, this is the core of legal indeterminacy: reasonable and well-informed people – and judicial decision-makers – can disagree about the interpretation and application of the law, especially when it is only broadly defined.

To be sure, in reference to indeterminacy we are not subscribing here to critical legal views of ‘strong’ legal indeterminacy, which would undermine the very legitimacy of the rule of law. Indeed, we are more concerned with what has been called ‘underdeterminacy’, that exists when ‘the set of results that can be squared with the legal materials is a nonidentical subset of the set of all imaginable results’ (Solum, 1987), as the analysis of national treatment seems to show. This is more in the vein of ‘moderate indeterminacy’ (Kress, 1989), which is inherent in law designed to address ‘hard’ cases, such as the distinction between protectionist measures and regulatory ones. In other words, we are by no means claiming that national treatment as a legal discipline is devoid of normative content, only that its determinacy is lacking.

This in itself is hardly shocking. The relative indeterminacy of the operative language of national treatment, under both Article III GATT and Article 2.1 TBT (and indeed, under other WTO agreements) is well recognized by WTO adjudicators. The famous ‘accordion’ analogy set forth by the AB in the *Japan–Alcoholic Beverages* Report is such an acknowledgement:

The concept of ‘likeness’ is a relative one that evokes the image of an accordion. The accordion of ‘likeness’ stretches and squeezes in different places as different provisions of the WTO Agreements are applied. The width of the accordion in

any one of those places must be determined by the particular provision in which the term 'like' is encountered as well as by the context and the circumstances that prevail in any given case to which that provision may apply.¹⁰

Notably, the 'accordion' analogy does not even tell us, as a general matter, in which direction 'likeness' stretches and squeezes in different provisions and different circumstances, when it should be broad and when it should be narrowly interpreted. We know how the AB used the analogy in *Japan–Alcoholic Beverages*, in creating a formal distinction between 'likeness' in Articles III:2 and III:4 GATT, but ultimately it did so on the basis of contextual interpretation which could have gone either way.

And in *EC–Asbestos* the AB commented even more frankly about the imprecision inherent in 'likeness', even within the same particular provision of the GATT:

[T]here is a spectrum of degrees of 'competitiveness' or 'substitutability' of products in the marketplace, and...it is difficult, if not impossible, in the abstract, to indicate precisely where on this spectrum the word 'like' in Article III:4 of the GATT 1994 falls.¹¹

We might conclude, therefore, that the judicial system, if not the Membership, has resigned itself to the indeterminacy of what is one of its core disciplines.

3.1.2 *National treatment as a standard*

One way of conceptualizing this indeterminacy in the national treatment discipline in the WTO is from the perspective of 'rules v. standards'. To WTO scholars and practitioners, this distinction will be most familiar from Joel Trachtman's classical discussion of the interaction between WTO law and non-WTO law (Trachtman, 1999), but we find this framework to be at least as illuminating in the current context. Rules are legal commands, like speed limits, that set out clear ex ante limitations on permissible conduct, leaving the adjudicator only with the task of applying the rule to the facts at hand. Standards are legal commands that are more fuzzy, requiring the adjudicator to establish both the applicable boundaries of permissible behaviour, and the facts. Standards often require balancing of divergent factors, leaving greater discretion in the hands of the adjudicator, ultimately leading to ex post determinations of permissibility. To Kaplow (1992), 'The only distinction between rules and standards is the extent to which efforts to give content to the law are undertaken before or after individuals act.' National treatment—both the product definition and the comparative 'less favourable treatment' element—is clearly a standard, not a rule. This has significant

¹⁰ See n. 7 above for details.

¹¹ EC-Asbestos (AB Report), para. 99.

implications for the development of the law and for its legitimacy. In Kaplow's economic analysis of rules versus standards, the pre-behavioural negotiation of rules has significant costs, whereas the post-behavioural application of standards may have significant legitimacy costs. This is certainly the case with respect to national treatment; despite the TBT, national treatment remains more of a standard than a rule, leaving most of the detailed specification to ex post analysis by Panels and the AB, whose legitimacy will always be costly to maintain.

One could envision two paths through which WTO Members might bring national treatment closer to being a rule. The first would be 'positive integration' in the form of agreed technical regulations; this exists, of course, to some extent, in Article 2.4 TBT's encouragement to use international standards as 'a basis for' national technical regulations. This is, however, subject to the flexibility allowed to Members when an international standard would be ineffective or inappropriate for fulfilling their legitimate objectives. And in many areas, international standards do not exist. National treatment (as well as the other requirements of the TBT) therefore persists as the governing discipline.

A second path would not engage with the technical regulations themselves but with the content of national treatment's operative terms, 'like products' and 'less favourable treatment', with detailed and concrete methodologies – similar to the way the Antidumping Agreement, for example, includes more detail on the application of economic concepts. In any case, for the time being, it will be up to Panels and the AB to apply national treatment (in TBT, GATT and elsewhere) as a 'standard', filling it with content through dispute settlement.

3.1.3 *Market analysis as a quest for determinacy?*

This seems straightforward enough, but now let us see how it applies in practice to the alternative judicial strategies employed by the Panel and AB in *US–Clove Cigarettes*, i.e., how they played different tunes on the same accordion. Two observations emerge. The first relates to the AB's transposition of national treatment in Article III to the TBT. This is the transposition of a standard and its application, by now generally accepted, relatively speaking, with respect to the GATT, and so the boat of judicial legitimacy is not exceptionally rocked. This is in contrast with the Panel's approach, which would have directly introduced to the likeness analysis questions of regulatory purpose, an interpretation of a standard in a way that would significantly alter the accepted practice regarding Article III:4 GATT. As we have seen, the AB's contextual interpretation that led to this conclusion is at least as reasonable as the Panel's.

More importantly, however, for present purposes, is the subsequent focus by the AB on competition-oriented analysis in the interpretation of product likeness, as an economic issue, *as if market analysis were more determinate*, and consequently more legitimate. The ostensibly less determinate regulatory policy distinctions are left to the second stage of 'least favourable treatment' analysis. It is on this judicial preference that we will focus the remainder of our commentary.

The AB's construction of Article 2.1 TBT seems to draw a line between purportedly objective economic factors that determine market relationships (taking into account factors that might also be the *reason* for a regulatory distinction, but can be observed in one of the traditional 'likeness' tests, as explained above), on one hand, and consideration of the policy purpose of a distinction between products (that have been found to be 'like'), on the other hand. Simplified, it is an attempt to distinguish between an economic 'fact' – 'competition-likeness' – and a political preference – regulatory distinction or 'policy-likeness' (using the terms of Mavroidis (2013) – meaning that the products are 'like' or 'unlike' for the purpose of the regulation). Analytically, this approach is not implausible and provides a degree of sequential clarity to an otherwise nebulous problem. It also serves to create what seems to be an objective limit to the responsibility of WTO Members to refrain from discrimination: the political bargain extends only to products in competition. However, we contend that to the extent that it is driven by a quest for greater normative *determinacy* it rests upon certain assumptions which are not entirely correct from an economic perspective. First, a competition-oriented approach to discrimination can provide important insights, but it does not attain the determinacy that the legal-normative sphere ostensibly lacks; economic market definition is itself dependent upon various functional and value-related determinations that are not very different from the legal tests. Second (as we will explain in greater detail in Section 5 below), an economic competition analysis would not so clearly defer the policy basis for a distinction between products that are otherwise 'like' to a separate stage of evaluation. The reasons for separating evaluation of 'market'-likeness from 'policy'-likeness are not to be found in economic analysis.

Interestingly, in (re)opening the question of determinacy in the law and economics of product definition, comparison, and the relative role of regulatory purpose, certain underlying similarities between the Panel and the AB emerge, perhaps explaining the relatively complacent critique by the AB of the Panel's approach. Both the Panel and the AB seem to believe that there is a bright-line distinction between 'market-likeness' that is based on objective (and presumably more legitimate) economic factors, on one hand, and a more subjective (and presumably less legitimate) 'policy-likeness'. The Panel interpreted the TBT as embedding the policy consideration into the economic analysis; the AB preferred to leave them apart. In the following paragraphs and sections, we endeavour to show why the differences between these two approaches, from an economic perspective, are much smaller than might have been assumed by both adjudicating instances, and that the determinacy offered by economic analysis does not solve the overarching problem of (moderate) legal indeterminacy inherent in national treatment.

Having framed the interpretative legal problems associated with product 'likeness', we turn now to a general exposition of indeterminacy in product definition from an economic perspective.

3.2 *Economic indeterminacy in product definition*

From an economic standpoint (which in itself is actually not very different from a judicial or legal perspective, except for its liberation from texts and rules of interpretation), one way to view the central issue of ‘likeness’ is a groping around for a decision rule that will let us readily determine whether two goods are like or unlike. With respect to Article 2.1 TBT, the Panel in *US–Clove Cigarettes* adopted one approach, relying on regulatory purpose. The AB put forward a different approach, relying heavily on consumer behaviour. How can one judge which approach is superior, from an economic perspective?

There are a number of features a good decision rule should have. It should hew tightly to its intended function. It should provide clarity to consumers, firms, and governments. It should strike an appropriate balance between the major competing interests at stake – frequently the desire to regulate in the public interest versus prevention of discriminatory treatment.¹² Ideally, the decision rule should be based on objective, observable factors, rather than subjective preferences. A key question is whether all of these desirable characteristics are consistent with each other.

As an attempt to sort out the results of this search for a decision rule, we introduce a modicum of formalism. The point of the exercise is to shed light on whether we can converge on an optimal decision rule; whether that decision rule would be objective or subjective; and to put some disparate points of argument in a common context. To foreshadow, we argue that a degree of subjectivity and value judgement will inevitably enter into the process. Even the economic analysis that might seem to offer the best hope of rigour and objectivity will have an inherently subjective component, and cannot cure the indeterminacy problem inherent in the ‘like product’ component of the national treatment norm.

3.2.1 *Describing the products*

In formal terms, each product X can be defined by N characteristics:

$$X^i = \{X_1^i, X_2^i, \dots, X_N^i\}$$

If $\forall i \neq j$ and $\forall k$ $X_k^i = X_k^j$ then the products i and j are identical in every characteristic and we say the products are identical (homogeneous). If there is any dimension k in which the products differ, then they are differentiated (heterogeneous).

As a way of organizing thoughts about the determination of whether products are ‘like’, this comprehensive approach lets us include an arbitrary level of detail about a product. We could have one set of characteristics that describe the physical

¹² See Howse and Levy (2012) for a discussion of the balance of interests in likeness determinations.

attributes of a product – its size, weight, packaging, chemical composition, flavour, etc. This is a particularly comfortable set of characteristics in that we would expect those characteristics to be readily observable and independent of any broader context.

We could imagine a second set of characteristics that relate to the way the product is produced or processed (processing and production methods or PPMs).¹³ Some PPMs would impact upon the physical attributes of the product (so-called ‘product related’ PPMs), such as organic vegetables compared to pesticide treated vegetables. Others, ‘non-product related’ PPMs – the wages and working conditions of the workers manufacturing the product, the environmental impact of the production process, etc. – would not have a readily observable impact on the product itself. This is a more controversial category of characteristics, but in our formal model there is no reason to a priori exclude them from the arbitrary degree of detail that distinguishes between products. It is where we could consider two apparently identical cans of tuna (i.e., products for which all of the first set of characteristics are equal) whose production methods differed in the threat they posed to dolphins (differences that would only show up in the second set of characteristics).¹⁴

We could consider a third set of characteristics that describe the way the product interacts with any given market. This could include the own-price elasticity of demand, the cross-price elasticity of demand with any number of other products, substitutability of supply, market size, the demographic profile of the product’s consumers (e.g., youthful newcomers to smoking versus long-time smokers). This set of characteristics should really prompt us to introduce an intertemporal component to our product definition, since there is no particular reason we should expect these characteristics to remain constant over time.¹⁵

There are yet other classes of characteristics, even farther removed from observability, that classify products not by physical characteristics or observable market statistics but by regulators’ purposes with regard to the product, consumer tastes and habits, or the end-uses to which the products may be put. We postpone a consideration of these issues for later, in Section 5 below. We turn now to the set of decision rules.

¹³ The literature on implications of PPMs for non-discrimination rules in international trade is vast; see Conrad (2011).

¹⁴ To be sure, it is this category in which questions of consumer information and hence labelling requirements enter into play. These issues did not arise in *US–Clove Cigarettes*, and we mention them here merely for the sake of completeness.

¹⁵ As already noted, this intertemporal problem arose in *US–Cloves Cigarettes (AB)*, but was not dealt with in substance. For present purposes, the intertemporal problem itself demonstrates indeterminacy in economic analysis, because there is no single rule defining what is the relevant period to be examined. In antidumping, by analogy, the Antidumping Agreement does not define the period to be investigated, and investigation authorities are relatively free to select it, although it must be specified in requests for information and in determinations.

3.2.2 Decision rules, unacceptable and otherwise

We now come to the true object of our interest – the decision rules. We define a decision rule L_i as a function of two variables,¹⁶ X_j and X_k , such that $L_i(X_j, X_k)$ takes on one of two values: {‘like’, ‘unlike’}.

The virtue of this comprehensive approach is that we can now consider the set of all potential decision rules, L_i . This is the set over which the ‘likeness’ jurisprudence has been searching, trying to argue that one rule is better than another. The previous subsection noted that we might usefully group characteristics into categories, such as product characteristics, descriptions of PPMs, and market characteristics – a list that was not meant to be exhaustive, although arguably it does cover most of the universe of conceivable differences between products, whether directly or indirectly. The relevance of those characteristics, and their relative weight in this construction, is entirely determined by the decision rule.

The process of searching for an ideal definition of ‘likeness’ can be thought of as one in which we take this large set of potential decision rules, L_i , and chisel it down to a subset with desirable features. For example, we can easily reject a decision rule that looks only at the characteristic of where the products are produced and declares two otherwise identical products to be ‘unlike’ solely because one is domestic and one is foreign. This, a type of rule often referred to as *de jure* discrimination, would allow for the most obvious sort of protectionism that the national treatment commitment is, at minimum, meant to preclude. A decision rule that referred primarily to the way two products were otherwise treated differently by law (e.g., different classification for tax or tariff purposes) would also not be acceptable, because it would essentially beg the question.¹⁷

This existence of some unacceptable decision rules establishes that some decision rules can be assessed as better than others. The question, then, is whether we can further refine the features of desirable decision rules to describe a set of superior approaches. This is effectively the exercise the GATT/WTO Panels have been undertaking for decades with respect to the GATT, and the Panel and AB were undertaking in *US–Clove Cigarettes* with respect to the TBT with the chief difference between the latter’s approaches being their emphasis on regulatory purpose as opposed to market competition. In the next sections, we consider the more observable facets of a product and their role in likeness determinations. Later sections take up questions of end-use, regulatory purpose, and consumer tastes and habits.

16 More precisely, it is a function of the vectors of characteristics that describe each of the products, allowing for comparisons of any of the elements.

17 Indeed, tariff classification is one of the ‘traditional’ tests of product likeness under Article III GATT, but it is of minor importance. In applying most-favoured nation treatment under Article I GATT, it is recognized that ‘fine product distinctions’ in tariff classification can be strategically used to discriminate between import sources (Hudec, 2000).

Physical characteristics. Physical characteristics are an obvious and necessary part of a desirable decision rule—and indeed they have always been part of the traditional ‘likeness’ tests in the GATT/WTO. One would not argue that a chocolate candy cigarette was the same as a clove or a menthol cigarette; quite simply, their chemical and physical composition is so different. The latter deliver nicotine when burned; the former do not. The latter can be smoked, but not eaten; the former are edible but cannot not be smoked.¹⁸ While this may be obvious, the question that quickly arises is whether physical distinctions offer any useful, objective guide for distinguishing between two products. It may be clear that edible and smokable cigarettes are different, but on physical characteristics, clove and menthol cigarettes are also different; they have different ingredients, different flavours and textures. But are these differences important?

If, as in the analysis of Howse and Levy (2012), we posit that the ideal decision rule L_i strikes the appropriate balance between permitting legitimate regulatory actions and precluding actions motivated by protectionism, the difficulty with decision rules based primarily on physical characteristics becomes apparent. A government bent on protecting domestic industry could readily justify any protectionist action it took simply by pointing out even minor—perhaps absurd—physical differences between two products.

To close this gaping loophole, it is necessary not simply to identify physical differences but to attach a judgement as to whether those physical differences are significant. As soon as we do so, however, we risk moving from objectivity to subjectivity. If we cannot identify, *ex ante*, which physical similarities are relevant and which are not, on the basis of a physical evaluation of the products alone, we will be unable to select among this particular subset of decision rules and will be unable to offer guidance to traders and Members as to which actions are permissible. This does not mean that physical characteristics will play a trivial role in the optimal decision rule. Rather, it suggests that, for clarity, we would like to find some readily observable indicator that offers guidance about the relevance of different physical characteristics.

Market characteristics. Attempting to distinguish like and unlike products simply on the basis of physical characteristics is clearly necessary but insufficient. This has led the jurisprudence, including the AB in *US–Clove Cigarettes*, to a consideration of market characteristics:

‘Likeness’ in Article 2.1 . . . is based on the competitive relationship between and among the product. (para. 156)

¹⁸ Notably, our examples are framed as physical characteristics even though they clearly relate also to ‘end-uses’, demonstrating that the AB was correct in finding that physical characteristics are not necessarily mutually exclusive from ‘end-uses’. The latter often derive from the former.

On its face, this might seem to offer just the objective indicator we are looking for. Perhaps two products could be ‘like’ if they target the same demographic in their marketing. Yet that particular indicator might lead us astray; we would not want to conclude that cigarettes and leather jackets are like products, just because they are sold to the same segment of the market. We thus wind our way around to the idea of product substitutability, which is indeed reflected in the concept of ‘Directly Competitive or Substitutable Products’ (DCS) in the *ad note* to Article III:2 GATT, that specifically applies to fiscal discrimination.¹⁹ This could lead to an adoption of economic Cross-Price Elasticity (CPE) analysis as (nearly) determinative of product definition and ‘likeness’.

This certainly seems to offer an advance beyond some of the difficulties plaguing other approaches. A decision rule that placed heavy emphasis on the demand CPE between two products is at least working off observable and measurable phenomena. Further, if the scrutiny under national treatment provisions is partially inspired by concern that imports might be illegitimately targeted to favour domestic industry, a CPE approach attempts to measure directly the extent to which restrictions on the imported good will have this effect, whether or not that was the regulator’s motivation.

However, to throw a wrench into the prospects of such an approach, while we may favour decision rules that place heavy weight on CPE variables, are we able to choose between them? Once we begin looking at cross-price elasticities, is there an obvious objective threshold beyond which we can say the two products are like? Or are we still left with subjective judgements, albeit better-informed ones?²⁰

Here an analogy to antitrust economics may prove useful. The objective of antitrust policy is different from national treatment. Antitrust is concerned with the degree of competition within a market and, *inter alia*, the potential price effects of allowing mergers or acquisitions. Thus, through antitrust law, the State regulates aspect of private activity. In contrast, national treatment is concerned with international competition, and through the norm (and dispute settlement), the WTO regulates state measures. Nevertheless, there are similarities in the types of data available and the central question of whether two products compete with each other correlates in an important way. The antitrust literature is also well developed,

¹⁹ See, e.g., Horn and Mavroidis (2004). Indeed the *Note ad* Article III:2 was the basis for the distinctions drawn by the AB in its earlier jurisprudence between the scope of the ‘accordion’ in Articles III:2 and III:4 respectively. The Panel in *US–Clove Cigarettes* also attached contextual weight to it, but the AB disagreed.

²⁰ Notably, while Panels and the AB have acknowledged the importance and usefulness of referring to econometric studies of CPE in assessing product definitions for the purpose of national treatment, they have also been careful to avoid establishing CPE as a definitive measure of ‘likeness’; see paras. 286–289 of the WTO Analytical Index, available at http://www.wto.org/english/res_e/booksp_e/analytic_index_e/gatt1994_02_e.htm#fnt413. We generally agree with this guarded approach, not only because product definition entails both quantitative and qualitative dimensions, but because the quantitative measures are themselves subject to indeterminacy.

having undergone decades of refinements. Indeed, some have suggested that economic methods of product definition employed in some antitrust jurisdictions should migrate to the WTO (Melischek, 2013).

In an antitrust context, the approach would seem enticingly simple: define the relevant market and then determine whether a proposed merger would result in an unacceptable degree of concentration in that market, in turn diminishing the degree of competition and raising prices. Despite the apparent simplicity and objectivity, Farrell and Shapiro (2010) describe a process plagued with subjectivity and arbitrariness:

The problems are particularly pronounced in the large class of mergers in which the merging firms sell differentiated products. . . Because of the differentiation, defining the relevant market can be problematic. . . When Amazon.com teamed up with Borders on-line, was the relevant market on-line book retailing or all book retailing? When Miller acquired Coors, was the relevant market domestic beer, all beer, all alcoholic beverages, or all beverages?

They note the centrality of market definition to antitrust analysis in the United States and conclude:

While much has been written in antitrust economics on how best to define markets, the fact is that in many differentiated-product industries, *there is no clearly right way to draw boundaries that are inevitably somewhat arbitrary.* (emphasis added – T.B., P.L.)

Kaplow (2010: 440) approaches the question from a legal perspective and is similarly despondent. He writes:

there does not exist any coherent way to choose a relevant market without first formulating one's best assessment of market power, whereas the entire rationale for the market definition process is to enable an inference about market power. Why ever define markets when the only sensible way to do so presumes an answer to the very question that the method is designed to address?

These are particularly daunting conclusions for the search for an ideal likeness decision rule, L_i . Arguably, the ultimate objective of antitrust efforts is clearer than those for national treatment.²¹ The data are at least as good. The economic methods have been tested and refined over time. Yet there is still the problem of subjectivity and arbitrariness.

Returning to the realm of national treatment and likeness, once we obtain a CPE of demand, what do we do with it? While it can certainly give a broad indication²²

²¹ In fact, the central thrust of the Farrell and Shapiro (2010) argument is to set aside secondary indicators of potential competition effects and adopt more direct predictors of what any given merger will do to the price. They are able to argue that their proposed decision rule is superior to existing ones because it hews more closely to the clear objective of antitrust policy.

²² We are actually being somewhat generous in our treatment of CPE as an indicator, setting aside potential concerns about ease of measurement and variability over time. Deeper concerns, such as whether

of whether there is a competitive relationship between two products, there is no clear bright line beyond which we obviously want our decision rule to signal ‘like’ and before which it should signal ‘unlike’

To sum up this section on economic determinacy, among the class of decision rules that rely upon market characteristics, we are still left with a subjective choice and without a clear winner. As a general conclusion, it would appear that a turn to economic analysis, and more specifically, the idea that focusing primarily on competitive factors in defining product ‘likeness’ will provide greater objectivity and determinacy through recourse to economic concepts, is exaggerated. At the very vague degree of detail and resolution provided for by the national treatment norms (whether in Article III GATT or Article 2.1 TBT), there remain numerous decisions to make in their application, even if a market approach is prioritized. And we have so far deliberately ignored the question of whether regulatory purpose cannot be taken into account as part of a competition-oriented analysis—a question we return to in Section 5 below. Now let us look in more detail at the correlations between legal and economic indeterminacy in one of the AB’s particular findings—relating to the ‘end-uses’ of cigarettes.

4. What are cigarettes for? Legal and economic indeterminacy and the AB’s ‘capability’ test for end-uses

4.1 The multiplicity of end-uses as a methodological problem

Beyond the array of characteristics already proposed as candidates for determining whether two products are ‘like’, the *US–Clove Cigarettes* Panel used a well-established test—both in GATT/WTO national treatment jurisprudence and in economic literature—of the ‘end-use’ of a product. When comparing menthol and clove cigarettes, the Panel found that they shared a common end-use: ‘to be smoked’. This was critical in the Panel’s determination that the two types of cigarettes were like products.

In its appeal, the United States argued that this was an excessively narrow approach (*US–Clove Cigarettes* (AB), para. 15). While menthol and clove cigarettes certainly could both be smoked, the United States argued that ‘end-uses’ are not just the use that is a ‘common denominator’ of products. There were other uses that were particular, in the US view, to each product, justifying a finding of ‘unlikeness’: menthol cigarettes ‘satisfy the nicotine addictions of millions of smokers in the United States’ while clove cigarettes are primarily used ‘for experimentation and special social settings’

certain competitive relationships can be observed at all, are deferred to the next section in which we address the ‘capability test’ relating to product end-uses.

As already noted above, the AB recognized this more detailed distinction between nicotine addiction and other end-uses, but on the facts found sufficient overlap among end-uses to preserve the finding of likeness. Importantly, that determination was based on more than just the way the two products are actually or principally used. The AB wrote:

The fact that more ‘addicts’ smoke menthol than clove cigarettes does not mean that clove cigarettes cannot be smoked to ‘satisfy an addiction to nicotine’ ... what matters in determining a product’s end-use is that a product is *capable* of performing it, not that such end-use represents the principal or the most common end-use of that product. (para 131)

From a generally critical legal perspective, this combination of findings has an internally contradictory effect. On one hand, the acceptance of more specific end-uses would seem to project the seriousness of the AB’s review – signalling to Panels that more detailed and rigorous investigation of end-uses is required; and more importantly, it would seem to increase the flexibility that Members have, in their capacity to distinguish between products that might otherwise be considered ‘like’. On the other hand, then comes the idea of a ‘capability’ test for end-uses, and much of this is washed away. A ‘capability’ test for end-uses necessarily expands the overlap between products, even *ad absurdum* (e.g., condoms are not only used for contraception and prevention of communication of diseases, but can also be used as water containers), thus *diminishing* Members’ flexibility in applying technical regulations to products that they might consider ‘unlike’. There is indeed no obvious legal or interpretative answer to the question and both the Panel’s and AB’s approaches are plausible (and in practice, led to the same outcome). Can economic analysis provide us with determinacy where legal determinacy is lacking?

From an economic standpoint, this issue actually raises a number of interesting – and parallel – questions. To what extent can we define a product by its end-use? How do we determine a product’s capabilities? Are those physical traits inherent to a product or do they depend on the ingenuity of the individuals in the target market? Do the uses change over time? Most relevant for the indeterminacy arguments of this paper: are there any clear objective criteria relating to end-uses that will help distinguish between like and unlike products?

4.2 *What is a product used for?*

There are at least two significant problems in determining the end-use of a product. First, the same act can have different components. Does smoking deliver nicotine? Yes. Does it give smokers something to do with their hands? Yes. Does it stimulate taste receptors in the brain? Yes. All of these are happening simultaneously. If we wish to argue that the two products are like, we will emphasize the overlap in these functions. If we wish to argue that the two products are different, we will emphasize the differences (e.g., perhaps clove and menthol cigarettes trigger different taste receptors).

One might be tempted to search for a threshold of overlap, such as saying that if X per cent of the potential uses are common, then the two products are like. The issue would then become the arbitrary breadth of descriptions. Just as some tariff classifications are very broad and others are finely detailed, we could describe the same activity broadly (the product is ‘smoked’) or narrowly (the product is held in the hands, it stimulates certain receptors, it delivers nicotine, etc.). Without some basic unit of measurement of usage, any comparison of overlap and adoption of a threshold will have an arbitrary component.

This indeterminacy carries over into the second major problem. Not only can the same usage be described in multiple ways, but there can also be multiple usages. An iPhone can be used as a flashlight. Computer dust cleaners can be used as drug delivery devices for a cheap high. (CBS Chicago, 2012). These are not the intended uses of the products, nor are they the primary uses of the products, but the products are capable of being used that way, establishing some degree of latent demand. This raises again the threshold question – how do we distinguish between primary and secondary uses? Would it be the number of people who are observed undertaking one activity versus the other? What happens if the same person uses the iPhone both as a communications device and to shed light? Aside from the near-impossibility of regularly obtaining usage data at such a fine level of detail, conceptually there would be no clear threshold for significance.

4.3 *What can this product really do?*

The previous section implicitly assumed that we would observe all the important uses of a product and would only have to grapple with the problem of how to assess the significance of that use. The ‘capability test’ adopted by the AB in *US–Clove Cigarettes*, as quoted above, makes the problem even more difficult. What if there are significant uses of a product that we *do not* observe under present circumstances, but that would emerge under different circumstances?

In fact, such a scenario emerges readily from a very simple economic specification. Suppose that we are focused on a certain subgroup of smokers who currently use menthol cigarettes and never smoke clove cigarettes. If we want to model their behaviour economically, we might start with a utility function. We could make the strong simplifying assumption that the utility they get from smoking is separable from all other components of their utility and that they have a fixed budget to devote to cigarettes.

Let M represent the number of packs of menthol cigarettes consumed (at price p_M each) and C represent the number of packs of clove cigarettes (at price p_C each). Let this sub-utility function take the simple form:

$$U(M, C) = 2M + C$$

Standard utility maximization will dictate that if $p_M < 2p_C$ consumers in this group will exclusively consume menthol cigarettes. If, under such circumstances, we asked which product served the end-use of pleasing this group, we would observe only

menthol cigarettes playing that role. We might be tempted to argue, based on that observation, that only menthol cigarettes delivered that sort of utility.

This highlights the importance of the capability test. By our construction of the utility function, we *know* the two products can serve the same end-use, yet we only *observe* one doing so. Nor is the capability to serve this end-use of merely theoretical interest. If the price of menthol cigarettes were to rise such that $p_M > 2p_C$ then this group would switch over completely to clove cigarettes to satisfy this end-use. Depending on the market structure in the cigarette industry, this capability of clove cigarettes could serve as a restraint on price increases by menthol cigarette producers, even if we are in the range such that no clove cigarettes are actually consumed. Thus, the unconsumed goods are in a competitive relationship with the consumed goods.

It is worth noting that this feature of demand in which consumers consume either one product or the other (except at the exact point $p_M = 2p_C$) is a feature of the particular form of the selected utility function. If, instead, we looked at a relatively common (non-linear) Cobb–Douglas utility function $U(M, C) = M^{2/3}C^{1/3}$ we would not get such ‘corner solutions’ in which one of the products goes unconsumed (note that utility is zero if M or C is zero). In this case, if a good could perform an end-use, we would expect to observe it.

But there is nothing wrong with the linear specification given earlier, nor is it unusual to think of products which people do not consume, but might under certain conditions, dependent on their potential end-uses. From an economic standpoint, the AB was correct to emphasize the capability of goods to meet end-uses rather than relying exclusively on observed behaviour. However, from the standpoint of achieving an objective means of discerning between like and unlike products, this capability test makes a difficult problem even harder, particularly if one posits that any product would have numerous, and indeed unrevealed, end purposes.

5. Where should regulatory purpose be considered? It doesn't really matter

5.1 *The role of regulatory purposes and distinctions*

In Section 3.2 above, we considered a range of product and market characteristics that may be used to determine whether or not two products are ‘like’. In so doing, we deliberately deferred discussion of some additional tests that have factored into national treatment jurisprudence in the GATT/WTO—the market-oriented characteristic of ‘consumer tastes and habits’ and the policy-oriented issue of regulatory purpose. It will be recalled that it is the relationship between these two factors that served as the basis for the most significant difference between the Panel and AB in *US–Clove Cigarettes*. The Panel incorporated regulatory purpose into its product ‘likeness’ analysis, specifically employing it to delimit the consumer group whose tastes and habits were assessed (youth and potential youth smokers).

In contrast, the AB reserved reference to regulatory purpose to the subsequent analysis of ‘less favourable treatment’, as a potential justification for detrimental effect on a product. We have also noted that there is no outstanding or obvious legal-interpretative reason to prefer one approach over the other – they can both be absorbed by the textual, contextual and indeed purposive dimensions of Article 2.1 TBT. Moreover, from an economic perspective, these two factors have in common a lesser degree of observability than simple physical product characteristics (which can be directly viewed or tested) and market characteristics such as CPE (which can be calculated from data, although subject to significant definitional flexibilities, as discussed above).

Importantly, to treat regulatory purpose on a basis fully comparable to other factors, we would not only have to include consideration of whether a product causes youthful users to become addicted to smoking (for example), but also of whether a particular regulator is convinced that the product has this effect, and perhaps, if deemed important, even whether this effect is a primary or secondary concern of that particular regulator. Some difficulties with this approach become quickly apparent: unlike market indicators such as CPE, it is not possible to observe the rank ordering of a particular regulator’s concerns (and politically, different regulators might carry different degrees of either clout or credibility). At minimum, the particular ordering we would be most concerned with is that of benign, legitimate concerns (e.g., health effects) and objectionable, protectionist concerns (e.g., well-being of the domestic industry). We could observe a regulator’s statements, including in legislative and preparatory documents about the rank ordering of his or her concerns, but these might not tell us the whole story. If some such statements about motivation will clearly be deemed WTO violations while others would be deemed acceptable, we should not be surprised to hear the pure, acceptable motivations advanced as overt justifications for regulatory action. Thus, there is an unsatisfactory susceptibility to manipulation that accompanies any decision rule that places heavy weight on unobservable intent. In this respect, the economic and legal perspectives and their problematics are essentially identical.

Indeed, cognizant of these and other problems, the jurisprudence on national treatment leading up to the *US–Clove Cigarettes* (AB) ruling featured advancing and then receding roles for regulatory purpose (or ‘aim and effect’ tests).²³ A certain ambivalence about regulatory purpose is laced through the AB decision as well. As described earlier, the AB contends that the determination of ‘likeness’ must be primarily based on the competitive relationship between and among the products. The AB decision then goes on to describe some of the difficulties of taking regulatory purpose into account: there may be multiple objectives at play and

²³ See Zhou (2012) for a discussion of the history of regulatory purpose in national treatment considerations, as well as an analysis of the *US–Clove Cigarettes* decisions.

Panels would be in a difficult position if they had to choose between them and assess their relative importance. This, the AB warns, could lead to 'somewhat arbitrary' results (para. 115).

Having made the case, the AB then adds (at para. 117):

Nevertheless, in concluding that the determination of likeness should not be based on the regulatory purposes of technical regulations, we are not suggesting that the regulatory concerns underlying technical regulations may not play a role in the determination of whether or not products are like.

This statement is to Article 2.1 TBT what *EC–Asbestos* (AB) was to national treatment under Article III GATT. It merely says that regulatory purpose can indirectly and spontaneously enter the 'like product' analysis insofar as it can persuasively be associated with a physical difference, an end-use or a consumer tastes trend.

The intertwined analytical questions raised by the use of regulatory purpose in national treatment considerations are of course not new. In the TBT context of *US–Clove Cigarettes*, but with implications for GATT Article III, we are now faced, however, with the question of whether it makes a tangible difference to consider regulatory purpose in product definition, or as an additional factor that might distinguish between products otherwise found to be like. In the next subsection we will return to first principles of economic analysis and put forward a simple model and then compare the various tests under set scenarios, dealing with this question. The subsequent subsection turns to the question of consumers' tastes and habits in the context of the role of regulatory purpose, which featured in the Panel decision and which raises some similar questions.

5.1.1 *A return to modelling*

A number of preliminary analytical questions here emerge. Is there any reason to think there is a single or dominant regulatory purpose? Is there any limit to the number of regulatory purposes that may apply? Does the regulator need to identify only a single legitimate purpose to justify an action? Or is there a threshold? Finally, does the inclusion of regulatory purpose seem likely to lead to a better likeness determination in any meaningful way?

Without promising answers to all of these questions, a simplified framework to think about the issue may be helpful. We return to the notation (employed in Section 4 above) of *C* denoting consumption of clove cigarettes and *M* denoting consumption of menthol cigarettes. Now, though, instead of taking the perspective of *consumer utility*, as in the previous discussion of end-uses, we consider the *utility of the regulator*. This may well differ from the well-being of consumers in some important ways. For example, suppose there is a negative externality to the consumption of the product. Almost by definition, the regulator might take this into account in a way individual consumers would not. Second, the regulator might benefit when a domestic industry flourishes. Thus, we will describe the regulator's

utility as:

$$U_R(b_c, -e_c, b_m, -e_m)$$

where:

- b_c represents the benefits of clove cigarettes
- b_m represents the benefits of menthol cigarettes
- e_c represents costs imposed by clove cigarettes
- e_m represents costs imposed by menthol cigarettes

We are being deliberately vague about the form this utility function takes and the relationship between the utility smokers receive through consumption (represented by b) and the utility gained by regulators. In general, if we think of a regulator as trying to promote social welfare, we will assume that consumer utility enters positively into regulator utility. We will also allow for the possibility that the regulator cares about factors that are not a concern to consumers. Externalities are one example of this (e). Another example might be additional weight attached to the domestic production of menthol cigarettes (M). This latter additional weight might emerge out of any number of political economy models. For our purposes here, we leave the specification general and make different restrictive assumptions in the sections that follow.

5.1.2 *The case of identical costs and benefits*

In an instance in which there is no qualitative difference between e_c and e_m , if they impose different costs on society, it will be only because one product is used more than the other. We can make the same extreme assumption for benefits. In this sense, we are beginning with an assumption of regulatory or ‘policy’ likeness. We will also simplify by limiting a regulator’s available actions. The regulator can either permit or ban cigarettes, excluding other regulatory measures. Thus, the possible strategies for the regulator maximizing U_R are four:

$$\{\text{Ban, Ban}\}\{\text{Ban, Permit}\}\{\text{Permit, Ban}\} \text{ or } \{\text{Permit, Permit}\}$$

We deem one strategy very unlikely right off the bat, though not impossible – banning the predominantly locally produced product while allowing the predominantly foreign produced one. If we allow for even minor political influence, however indirectly, the benefits of the domestically produced good (M) are likely to receive a higher weight in U_R than the benefits of the imported good (C). If the regulator is to ban one product, it will be the one whose benefits receive less weight in that welfare analysis – the imported product, C .

Thus, we are left with two pooling strategies (both products treated the same – either banned or allowed) and one separating strategy (a distinction between the products, banning the foreign and allowing the domestic). If we assume that the negative externalities are sufficiently serious that the regulator requires there to be some abatement, then we are down to two possibilities: banning the import, or banning both.

With the strong assumptions of this scenario, in the cloves–menthol dyad, the regulator would initially like to ban *C* and permit *M*. If the public perceives them as perfect substitutes and the regulator attaches greater weight to the domestically produced *M*, then blocking imports and allowing *M* to take over the market would leave the total level of benefits ($b_c + b_m$) unchanged, but raise U_R . By assumption, consumers do not care which variety of cigarette they smoke; regulators do care because they derive benefits from local production. Thus, a policy that tilts consumption away from *C* and towards *M* would only help the regulator. The regulator, however, bound by national treatment restrictions, would find this policy difficult to justify. If *M* replaces *C*, and they have identical externalities, then the overall costs are left unchanged as well ($e_c + e_m$). The only change was in regulator welfare, which would not count as legitimate.

The regulator, in this scenario, would try to argue that the products are not, in fact, 'like'. They would differ in some dimension – a physical characteristic, for example – that could be used as justification for the distinction. As we argued earlier, it is difficult to come up with any bright line that separates, in the abstract, legitimate from illegitimate distinctions in this respect. That is one reason there is a temptation to examine regulatory purpose. In our model, we have only allowed the regulator one legitimate regulatory purpose: diminishing negative externalities (one might think of these as well-established health effects, for example). We have constructed the case so that this legitimate regulatory purpose does not justify differential treatment. But the regulator could claim another purpose – protecting the public from some additional alleged harm e'_c caused by smoking clove cigarettes. The regulator certainly has the incentive to make such a claim. Otherwise, the distinction is almost certainly condemnable as simply related to the origin of the products, and hence, protectionist and discriminatory.

Here we see a potential difference in result between the various tests. A physical characteristics test could draw a rather arbitrary distinction between the two products. A competitive relationship test could not provide a distinction (under the strong assumption of perfect substitutes). A 'regulatory purpose' test potentially could find a distinction, with the arbitrariness or legitimacy of that distinction a key question. If the concern about e'_c were deemed to be a legitimate regulatory purpose, then the discriminatory policy would have good cause to be upheld. Alternatively, it could be the rejection of the concern about e'_c that prevents the claim about unlike products (based on physical characteristics) from going through.

To be sure, our model was constructed so there was a clear, right answer, if all is known. Yet, notably, the inclusion of regulatory purpose as a determining factor did not guarantee that the right answer would be reached. Furthermore, there was a clear right answer here because of factors that are unobservable in reality: the true objective functions, the actual negative effects. In practice, there would be difficult policy calls on which concerns are legitimate and which are illegitimate – exactly the point the AB addressed in *US–Clove Cigarettes* when it worried about Panels

sitting in judgement and reaching potentially arbitrary conclusions in product definition.

Importantly for the question of ‘sequencing’ between consideration of regulatory purpose as a factor in product definition as opposed to ‘less favourable treatment’, this model demonstrates that there is not much daylight, if any, between the difficult choice based on physical characteristics and the difficult choice based on the legitimacy of the regulatory purpose. The question ‘do the differences in physical make-up matter?’ is in many tough cases essentially the same as the question ‘is the regulatory concern about e'_c legitimate?’. While regulatory purpose may not offer a clear, easy resolution, it is difficult to see how it can be completely excluded from the discussion, given the indeterminacy of other characteristics. Consequently, in this example, it is – from an economic perspective – very difficult to disentangle the ‘likeness’ determination stage from the ‘less favourable treatment’ stage. Following AB guidance, one could conduct an initial likeness determination, relying on competitive relationships, and then determine that less favourable treatment was justified by a legitimate regulatory purpose. Or one could determine that the physical differences made the products unlike, based upon a determination that the regulatory purpose supported that focus on the products’ physical make-up (i.e., concern about e'_c is legitimate). Each route seems to arrive at the same place.

5.1.3 *Different benefits, same costs*

We can now begin to loosen the strong assumptions of the previous section by allowing the ‘benefits’ of the different types of cigarettes – or any other product, for that matter – to be qualitatively different. We can assume they are quantitatively similar, but clove cigarettes deliver that level of utility to one group of consumers and menthol cigarettes deliver that level of utility to an entirely separate group of consumers. Tastes are assumed to be such that neither group of consumers derives any utility from consuming the other product. We maintain the assumption that negative externalities are the same for each product.²⁴

There are two immediate differences that come into play. First, the products are no longer in a competitive relationship. Under these assumptions, movements in the price of clove cigarettes would have no effect on demand for menthol cigarettes. Thus, the cigarettes would be ‘unlike’ according to that portion of the test (which may indeed be dispositive – in the absence of a competitive relationship, there is little cause for trade law to intervene on the basis of discrimination). Second, in contrast to the previous example, a discriminating strategy by the regulator {Permit, Ban} would reduce the negative externality problem. If clove

²⁴ We noted that it was unlikely that the domestic product would be banned while the import permitted, but technically, even if this happened, this would not provoke WTO scrutiny. Of course, more complicated scenarios involving third parties and MFN issues might ensue.

cigarettes were banned, e_c would disappear and would not be replaced by additional levels of e_m .

We still leave our regulator with one legitimate purpose—the abatement of e . We still assume that the products are identical in their respective production of e . The pay-offs to the regulator are somewhat different, since the fall-off in benefits from b_c will not be replaced by new b_m benefits. Whereas consumers—about whom the regulator does care—used to be pacified by a switch from C to M, under this set of assumptions, that option will offer no solace. In this case, the regulator chooses between (full abatement, no loss of benefits), (partial abatement, some loss of benefits), and (no abatement, full loss of benefits). Suppose, for a moment, that the consumption levels of clove and menthol cigarettes are identical.

Under these assumptions, there are physical differences between clove and menthol cigarettes. Furthermore, the cigarettes are not in a competitive relationship. On these bases alone, the products would not be considered 'like'. However, for the legitimate regulatory purpose of abatement of e , they are identical. Again, if abatement is sufficiently important, the regulator could opt to ban imported clove cigarettes, but not menthol. As constructed, the sole motivation for not also banning menthol cigarettes would be the additional weight placed by a ban on domestic industry, since all other factors were assumed equal. This would seem to be exactly the sort of discrimination that national treatment was intended to protect against, but in the absence of a competitive relationship between the products, national treatment would probably not apply. The regulator could justify the action on the basis of physical characteristics (though there would be the persistent question of which are important), and refer to CPE analysis that shows that the products are not in a competitive relationship. The regulator could also claim concern about the additional e_c' , as before, but given the physical differences and lack of competitive relationship, this would be superfluous. Importantly, for present purposes, it would not make a difference at what stage of the analysis the question of regulatory purpose were considered. If it were taken as part of the product definition, it could not overcome the physical differences and absence of competitive relationship. If it were deferred to the analysis of 'less favourable treatment', it would not even be considered, because the products would not be found to be 'like'. Comparing the treatment accorded to groups of products that are not in competition with each other would significantly overstretch the political bargain embodied in the national treatment discipline in the GATT and the TBT, and such a measure would more properly be examined as a quantitative restriction under Article XI GATT, to be justified (or not) under Article XX GATT.

5.1.4 *Different benefits, different costs*

Finally, we can allow the negative externalities of the products in question to differ as well. This could occur in any number of different ways. Suppose that it is a question of magnitude; each cigarette poses health risks, but $e_c = \alpha * e_m$, where $\alpha > 1$.

In this case, the regulator may legitimately wish to ban the more damaging product and a regulatory purpose test could uphold the action (even if, as in the previous section, a competitive relationship test did not). There would still be a decision to be made about how big α would have to be to justify this differential treatment.

Overall, our model has shown that across what is arguably the entire spectrum of relationships between products in terms of regulator utility derived from permitting or banning their sale on the market, it does not matter whether the regulatory purpose of the distinction is considered as part and parcel of the definition of ‘like products’, or as a separate justification negating the existence of ‘less favourable treatment’. The results are, in any event, and subject to the various tricky judgement calls that need to be made, the same. As we have already noted, however, there are advantages to separating the assessment of regulatory purpose from product definition, namely, the relative clarity (if not determinacy) the separation provides, the transparency gained from a separate examination of regulatory considerations, and the comfort and legitimation provided to Members, that non-discrimination is required only when the products are in a competitive relationship.

5.2 *Which consumers’ tastes and habits?*

One of the differences between the *US–Clove Cigarettes* Panel and AB decisions was the role allotted for consumer tastes and habits. In particular, did it make sense to distinguish youth from adult smokers? This approach – certainly in the Panel’s view – links regulatory purpose to market definition. If the purpose is to reduce youth smoking, the Panel found that it could focus its inquiry on the tastes and habits of youth smokers, rather than all smokers. The AB disputed this but found ‘it is not necessary to demonstrate that the products are substitutable for all consumers or that they actually compete in the entire market. Rather, if the products are highly substitutable for some consumers but not for others, this may also support a finding that the products are like’ (para. 142). Thus, ‘while the Panel should not have limited its analysis of consumer tastes and habits to young and potential young smokers to the exclusion of current adult smokers, this does not undermine the Panel’s finding. This is because the degree of competition and substitutability that the Panel found for young and potential young smokers is sufficiently high to support a finding of likeness under Article 2.1 of the TBT Agreement’ (para. 145).

A number of the analytical points from earlier in this article are relevant here. From Section 4, there is the finding that even under well-behaved preferences, the fact that a subpopulation of consumers does not consume one variant of a product need not mean that they will never consume that product. Further, it is still possible for the consumed and unconsumed variants to be in a competitive relationship. From Section 5.1, all the problems of regulatory purpose apply. We presumably only end up looking at the tastes and habits of a subpopulation in a national treatment case when those tastes and habits are sufficiently different from those of other portions of the general consuming population to support an argument for

differential policy treatment. In such cases, we are left with the fundamental problems of regulatory purpose: true purposes are unobservable and there will likely be ample room for regulators to describe new purposes that serve to differentiate the products in question; and of course, the underlying policy question remains – how to differentiate between legitimate and illegitimate purposes?

5.3 *Political space and political costs*

Whether in looking at the legitimacy of regulatory purpose in defining like products or in determining less favourable treatment, a principal concern raised by critics of alleged WTO overreach is the restriction of policy space. Perhaps the principal cost of an overly expansive interpretation of whether two products are 'like' is to limit the range of permissible actions taken by the government adopting the discriminatory measure.

The illustrative examples above provide a relatively precise way to think about the limitation of policy space. That policy space was limited to four particular options: {Ban, Ban} {Ban, Permit} {Permit, Ban} or {Permit, Permit}. Following the approach of that section, let us assume, *arguendo*, that the products have the same benefits and negative externalities; they are, in fact, like. In this case, if the national treatment clause is functioning well, the policy space is reduced from four options to three. If the legitimate regulatory purpose of reducing the negative externalities is the true goal of the regulator, there still remain policy choices that would meet this goal.

If this scenario accurately described the *US–Clove Cigarettes* case – and many of the assumptions made here would be disputed in practice – revealed preference would indicate that the restriction of policy space was not innocuous. US policy-makers selected the impermissible option, the one that banned imported clove cigarettes but allowed consumption of domestically produced menthol cigarettes to proceed. The US side found little solace in the argument that it was permissible to ban *both* clove and menthol cigarettes because it argued that the banning of menthol cigarettes would be too costly, while the evidence produced by Indonesia suggested that the exclusion of menthol was the result of protectionist pressures. That is a policy judgement that the US would have to make.²⁵

We would note, though, that if the national treatment clause is to mean anything, there must be instances in which it restricts policy space and imposes additional costs. As in the simple example above, it will generally be the case that domestic producer interests are better represented in domestic regulatory deliberations than foreign producer interests. This will provide a persistent incentive, when searching for a way to meet regulatory interests at minimal cost, to impose those costs on

²⁵ Indeed, the US recently announced its intention to comply with the AB Report through steps taken with a view to increased regulation of menthol cigarettes.

foreign exporters rather than domestic producers. By requiring no less favourable treatment, the national treatment clause moves in the direction of putting foreign producers on an equal footing in the eyes of regulators. It should be no surprise that having to worry about the welfare of all groups rather than being able to exclude one group should prove more costly.

Moreover, in keeping with this article's general argument, we find no particular difference, from an economic perspective, between a consideration of regulatory purpose as either a part of product definition, or as a separate factor justifying 'less favourable treatment'. If anything, there are reasons to separately review regulatory distinctions, both economic and legal, for reasons of transparency and legitimacy. Thus, overall, we find merit in the AB's approach.²⁶

6. Conclusions

In both law and economics there is a desire to provide well-defined, objective answers to key policy problems. Aside from the eagerness to settle on optimal policy, clear answers—based on data and straightforward rules or precedents—provide an essential predictability to governance and let businesses and legislators avoid costly litigation, uncertainty and delays. The *US–Clove Cigarettes* case represents an advanced episode in the ongoing struggle to clarify the question of what it means for two products to be like for the purpose of establishing the existence of trade-distorting discrimination in the context of national treatment. While the AB, like the Panel before it, found in favour of the complainant, it employed different reasoning from that used by the Panel. That resolved the case—at least at the first level of litigation—and clarified the meaning of national treatment in Article 2.1 TBT, but nevertheless left room for doubts about where lines would be drawn in future cases.

In this article, we have discussed the legal and economic reasons that such indeterminacy persists. Some of the challenge comes out of the leap from GATT Article III jurisprudence to the application in Article 2.1 TBT. The language of the two sections allows room for dispute about whether they should have identical or distinct interpretations. From an economic standpoint, we discuss how none of the prominent, readily observable candidates for determining the likeness of products—whether physical or market characteristics—is sufficient for distinguishing between meaningful and spurious distinctions. As soon as one introduces—inevitably—the concept of regulatory purpose, which can help distinguish acceptable from unacceptable actions, one moves well into the realm of unobservables. We thus conclude that a degree of indeterminacy is inherent in the determination of likeness,

²⁶ Thus, we differ not only with the *US–Clove Cigarettes* Panel, but with other academic voices such as Mavroidis (2013) and Mavroidis (2014) forthcoming in this volume, as well as Regan (2006: 195) ('if we understand that we must consider regulatory purpose somewhere, the "likeness" inquiry seems the most natural place').

not just a temporary condition as the jurisprudence is refined. Importantly, this indeterminacy is shared both in legal and economic analyses. Indeterminacies in the law that occur when normative guidance is limited to standards rather than rules cannot be cured by depicting the issues as of a primarily economic nature, or by mere recourse to quantitative economic tests.

Accepting the importance of regulatory purpose still leaves the question of the analytical stage at which this factor should be considered: as part of product definition, or as a factor in the comparison of treatment accorded to products found to be otherwise 'like'. Indeed, this was one of the most significant differences between the Panel and AB analyses in *US–Clove Cigarettes*, and a prime example of legal indeterminacy in national treatment. To address this, we modelled a 'best case scenario' in which an omniscient arbiter could easily see the 'right' answer to the question of whether a measure was protectionist or not. We showed that, in the absence of such omniscience, the decision would still require difficult judgement calls.

Moreover, we found that from an economic perspective, it makes no difference at which stage those calls are made – in defining the product range, or in comparing the treatment of the products. However, dealing with regulatory distinctions between products as a separate stage, after product definition, has advantages in preserving the legitimacy of the process: it separates regulatory purpose from more objectively observable factors, preserves the political bargain whereby differential treatment is problematic only when products are in a competitive relationship, and sustains the transparency of policy considerations.

Although our discussion has been based on a particular case adjudicated under the TBT agreement, the analysis could just have well dealt entirely with national treatment under Article III GATT. Indeed, this indicates that the AB was right in anchoring its approach to Article 2.1 TBT in the GATT; even more importantly, this suggests that the AB's new jurisprudence on product definition and regulatory purpose should carry-back into the GATT itself, where applicable.

The indeterminacy that plagues the issue of product likeness does not imply that the issue is hopeless. It simply highlights that there are some tasks better performed by a legislative body than a judicial body. The distinction between permissible and impermissible regulatory purposes and acts depends heavily on value judgements. A new political agreement among WTO members could, in theory, provide clarity on these judgements. If the issue is left to judicial bodies without further guidance, however, the *US–Clove Cigarettes* case is unlikely to be the last one grappling with the issue of product likeness.

References

- CBS Chicago (2012), "'Dusting' High a Deadly Fad for Young People", at <http://chicago.cbslocal.com/2012/11/12/dusting-high-a-deadly-fad-for-young-people/>, 12 November.
- Choi, Won-Mog (2003), *'Like Products' in International Trade Law: Towards a Consistent GATT/WTO Jurisprudence*, Oxford: Oxford University Press.

- Conrad, C. R. (2011), *Process and Production Methods (PPMs) in WTO Law: Interfacing Trade and Social Goals*, Cambridge: Cambridge University Press.
- Farrell, J. and C. Shapiro (2010), 'Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition', *B.E. Journal of Theoretical Economics*, 10(1) (March).
- Flett, J. (2013), 'WTO Space for National Regulation: Requiem for a Diagonal Test Vector', *Journal of International Economic Law*, 16(1): 37–90.
- Horn, H. and P. C. Mavroidis (2004), 'Still Hazy after All These Years: The Interpretation of National Treatment in the GATT/WTO Case-Law on Tax Discrimination', *European Journal of International Law*, 15: 39–69.
- Howse, R. and P. Levy (2012), 'The TBT Panels: *US–Clove Cigarettes*, *US–Tuna*, and *US–COOL*', in C. P. Bown and P. C. Mavroidis (eds.), *The WTO Case Law of 2011*, American Law Institute Reporters' Studies, Cambridge: Cambridge University Press.
- Hudec, R. E. (1998), 'GATT/WTO Constraints on National Regulation: A Requiem for an "Aim and Effects" Test', *International Lawyer*, 32: 619–649.
- (2000), '"Like Product": The Differences in Meaning in GATT Articles I and III', in Thomas Cottier and Petros C. Mavroidis (eds.), *Regulatory Barriers and the Principle of Non-Discrimination*, Ann Arbor: University of Michigan Press, pp. 101–123.
- Kaplow, L. (1992), 'Rules versus Standards: An Economic Analysis', *Duke Law Journal*, 42: 557–629.
- (2010), 'Why (Ever) Define Markets?', *Harvard Law Review*, 124: 437.
- Kress, K. (1989), 'Legal Indeterminacy', *California Law Review*, 77: 283.
- Marceau, G. (2013), 'The New TBT Jurisprudence in *US–Clove Cigarettes*, *US–Tuna II* and *US–Cool*', *Asian Journal of WTO and International Health Law and Policy*, 8(1): 1–39.
- Mavroidis, P. C. (2013), 'Driftin' Too Far from the Shore – Why the Test for Compliance with the TBT Agreement Developed by the WTO Appellate Body Is Wrong, and What Should the AB Have Done Instead', *World Trade Review*, 12(3): 509–531.
- Melischek, C. A. (2013), *The Relevant Market in International Economic Law: A Comparative Antitrust and GATT Analysis*, Cambridge: Cambridge University Press.
- Regan, D. H. (2006), 'Regulatory Purpose and "Like Products" in Article III:4 of the GATT (with Additional Remarks on Article III:2)', in G. A. Bermann and P. C. Mavroidis (eds.), *Trade and Human Health and Safety*, Columbia Studies in WTO Law and Policy, Cambridge: Cambridge University Press, pp. 190–223.
- Solum, L. B. (1987), 'On the Indeterminacy Crisis: Critiquing Critical Dogma', *University of Chicago Law Review*, 54: 462.
- Trachtman, J. P. (1999), 'The Domain of WTO Dispute Settlement', *Harvard International Law Journal*, 40: 333.
- Zhou, W. (2012), '*US–Clove Cigarettes* and *US–Tuna II (Mexico)*: Implications for the Role of Regulatory Purpose under Article III:4 of the GATT', *Journal of International Economic Law*, 15(4): 1075–1022.