

Communication and Interpretation Challenges Related to Economic Evidence in Trade Disputes^a

Marios Iacovides and Marion Jansen^b

Abstract

Recent years have witnessed numerous instances in which economic evidence has been submitted to adjudicators in the context of WTO disputes. As it turns out, adjudicators have used this evidence only hesitantly as a basis for their decisions. In this paper we argue that a number of communication and interpretation challenges arising from the use of quantitative economic evidence can explain this phenomenon. In particular, we argue that it is in the current context difficult for adjudicators to assess the reliability and (un)biasedness of such evidence. Guidelines on how to assess quantitative evidence and benchmarks against which to evaluate the quality of such evidence may represent a useful if not necessary step in order to raise the acceptance of the use of quantitative evidence in trade disputes.

^a This chapter was partly authored with support from the project ‘Dispute Settlement in Trade: Training in Law and Economics’ (www.dissettle.org), a Marie Curie Initial Training Network funded under the EU Seventh Framework Programme (Grant Agreement No FP7-PEOPLE-2010-ITN_264633). The chapter will appear in Jansen, Pauwelyn and Carpenter (Eds.): “The Use of Economics in International Trade and Investment Disputes”, Cambridge University press. The authors would like to thank A. van Aaken, R. Teh, and H. Yano for useful comments on previous versions of this paper. All errors and omissions remain the authors’ and views expressed are personal and do not reflect the views of the institutions they are affiliated to.

^b Marion Jansen (jansen@intracen.org) is Chief Economist at the International Trade Centre, Geneva, Switzerland. Marios Iacovides (marios.iacovides@kkv.se) is Case Officer at the Swedish Competition Authority and a postdoctoral fellow in EU law (competition law and WTO law) at Uppsala University, Sweden.

© The Author.

All rights reserved. No part of this publication may be reproduced without the permission of the author.

Centre for Trade and Economic Integration (CTEI)

The Centre for Trade and Economic Integration fosters world-class multidisciplinary scholarship aimed at developing solutions to problems facing the international trade system and economic integration more generally. It works in association with public sector and private sector actors, giving special prominence to Geneva-based International Organisations such as the WTO and UNCTAD. The Centre also bridges gaps between the scholarly and policymaking communities through outreach and training activities in Geneva.

www.graduateinstitute.ch/ctei

1. Introduction

Recent years have witnessed numerous instances in which legal and economic experts were expected to work together in the context of trade disputes. World Trade Organization (WTO) Members' submissions increasingly contain economic content, for instance, in the form of references to econometric or simulation evidence. In such cases, WTO panels tend to seek economic guidance on how to handle the evidence, notably through the involvement of economic experts in the WTO Secretariat teams that assist them. In practice, at least two types of situations arise in which WTO adjudicators cannot avoid dealing with economic analysis:¹

- (i) In disputes –and related appeals - where at least one of the parties submits economic evidence, for instance, in the form of a quantitative economic model; and
- (ii) At the arbitration stage of proceedings, where the determination of countermeasures may require a quantitative estimate of trade effects.

Dealing with economic evidence in WTO disputes has proven to be a challenge. Numerous panel reports reflect a certain unease of panellists to constructively use economic arguments. Quantitative evidence has frequently been ignored or dismissed by panellists. This attitude stands in stark contrast with the high popularity of quantitative economic evidence among policy makers, including those active in the area of trade. It is common for simulation evidence to be used in high level policy debates when it comes to providing arguments in favour of or against certain types of policies. For instance, statements on the gains from trade policy changes in terms of GDP, employment, or wages, are typically based on simulation methods. They easily make their way into media headlines and are frequently found in policy documents produced by or upon request of national governments and international organizations.² The hesitance of adjudicators to use economic evidence in trade disputes also stands in contrast with the situation in the domain of competition policy where it has become standard for adjudicators to use and rely on economic evidence.

This situation has given rise to a rich debate among academics, practitioners, and stakeholders on exactly what role economics should play in the interpretation and application of WTO law.³ In this chapter, we want to contribute to this debate by asking the

¹ See World Trade Organization, *World Trade Report 2005: Exploring the links between trade, standards and the WTO*, (2005), available at http://www.wto.org/english/res_e/publications_e/wtr05_e.htm (last accessed 25 September 2014).

² See, for instance, various sustainability impact assessments of regional trade negotiations produced on behalf of the European Commission: <http://ec.europa.eu/trade/policy/policy-making/analysis/sustainability-impact-assessments/>; evaluations by NGOs or international institutions of the economic impact of a conclusion of the Doha Round (e.g. S. Polaski, *Winners and Losers: impact of the Doha Round on developing countries*, (2006) available at <http://carnegieendowment.org/files/Winners.Losers.final2.pdf> (last accessed 19 September 2014); K. Anderson, W. Martin and D. Van der Mensbrugge, 'Would multilateral trade reform benefit Sub-Saharan Africa?', *World Bank Policy Research Working Paper* 3616, (World Bank, 2005), available at <http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-3616> (last accessed 19 September 2014); or a report to the G20 on the employment effects of trade liberalization OECD, ILO, World Bank and WTO, *Seizing the Gains from Trade for Employment and Growth*, (2010), available at <http://www.oecd.org/trade/benefitlib/46353240.pdf> (last accessed 19 September 2014).

³ See, for instance, on the interpretation of 'likeness' and 'substitutability': W.-M. Choi, *'Like Products' in International Trade Law: Towards a Consistent GATT/WTO Jurisprudence* (Oxford University Press, 2003); A. Emch, 'Same Same But Different? Fiscal Discrimination in WTO Law and EU Law: What Are 'Like'

question why legal experts may find it hard to deal with economic evidence in the context of trade disputes. We will argue that there are numerous reasons for this and that several of them are linked to problems of communication between economic and legal experts. This in turn makes it difficult for adjudicators to interpret economic evidence in the context of trade disputes and as a result the evidence is rarely used as a basis of judgements.

Significant differences in the working methods between economists and lawyers can go a long way in explaining why economists and lawyers may find it difficult to talk to and work with each other in the field of international trade. We describe these differences in section 2, where we notably emphasize the duality between the role of 'specificity' in the field of economics and the role of 'general applicability' and 'ambiguity' in law. In section 3, we illustrate how communication challenges arising from this duality can affect jurisprudence. This is done by providing a short digest of WTO disputes in which economic evidence has been submitted.

In section 4 we explain why the challenges described in previous sections do not impede policy makers from heavily using economic evidence in trade negotiations and communications about trade and why they do not impede adjudicators from using economic evidence at the arbitration phase. We end this chapter by proposing venues for facilitating the use of quantitative economic evidence at the trade dispute stage and by arguing in favour of an increased use – by economists and lawyers – of general economic concepts (as opposed to data analysis) in the context of trade disputes.

2. Overcoming Differences in Working Methods: an Uphill Struggle

Working methods have a profound impact on individuals' ways of communicating on their fields of expertise. In the case of economists and lawyers this can lead to situations where lawyers are flabbergasted by the response received from economists to real life questions related to disputes.

We start this section with a number of examples of such situations of –sometimes profound - misunderstandings before delving deeper into the underlying reasons for these misunderstandings in the following subsections.

The first example refers to difficulties resulting from the fact that economists often rely on mathematical or statistical terminologies when explaining economic phenomena. International economics as it has been taught in recent decades is strongly influenced by the neoclassical school and is heavily based on the use of mathematics and statistics. Economists will therefore tend to think in terms of numbers, variables, graphs and equations, not in terms of argumentation as used in law. In answer to a question by a lawyer regarding the effect of a domestic policy on prices or trade flows, for instance, an economist may start drawing a graph and highlight the interceptions of a number of lines; an answer that is likely to leave the lawyers puzzled.

Products?', (2005) 32:4 *Legal Issues of Economic Integration* 369; J. B. Goco, 'Non-Discrimination, "Likeness", and Market Definition in World Trade Organization Jurisprudence', (2006) 40:2 *Journal of World Trade* 315; N. Diebold, *Non-Discrimination in International Trade in Services: 'Likeness' in WTO/GATS* (Cambridge University Press, 2010); C. A. Melischek, *The Relevant Market in International Economic Law: A Comparative Antitrust and GATT Analysis* (Cambridge University Press, 2013); and cf. S. B. Wille, 'Recapturing a Lost Opportunity: Article III:2 GATT 1994 *Japan-Taxes on Alcoholic Beverages* 1996', (1998) 9 *European Journal of International Law* 182.

The second example refers to frustrations resulting from the fact that economists rarely rely for their answers on general economic arguments. In order to fit complex human interactions into mathematical or statistical models, economists have to make numerous assumptions. When assumptions change, the findings of models change. When asking an economist “what happens if” the lawyer will therefore often receive the answer “it depends”, which may only be mildly helpful for the lawyer who needs to take a clear-cut decision.

The third example refers to the irritation that can arise when economists fail to accept that legal texts like the GATT Agreement are supposed to be the starting point of their analysis. While standard text book references exist for economists, academic economists will spend much of their working life developing new approaches and doing things ‘differently than before’. The idea of continuity does not play much of a role in modern economic thinking, in particular in academia. When discussing the application of a provision in a trade agreement, economists – notably junior economists - may therefore not hesitate to argue that the legal text is wrong and needs to be changed. Legal experts who are familiar with the challenges of law making are likely to find such remarks rather unconstructive and naive.

A closer look at the underpinnings of misunderstandings arising from different working methods can be useful for finding ways to overcome them.

a) Talking to each other: the precision of language versus the precision of mathematics

As mentioned before, mathematical and statistical tools play an important role in International economics as it has been taught in the 1990s and early 2000s. Mathematical formulae represent very precise tools to analyse specific economic questions. Typically, when economists communicate about these questions with their peers they do so by making reference to those mathematical formulae. Therefore, communication takes the form of a combination of language and reference to abstract parameters used in the relevant formula. Economic terms like “equilibrium price” or “first best policy” tend to gain their meaning through reference to a mathematical formula; that is to say, a word corresponds to a mathematical definition. The same word may have a different meaning when referring to a different formula. Ultimately, the internal logic of an argument is achieved through a combined use of wording and mathematical formulae.

When working with legal experts not used to applying formulae, the economist needs to be able to build arguments using verbal communication only, that is, without reference to formulae or parameters. Economists are not necessarily used to doing this as the skill of verbal communication is not necessarily emphasized during their education. This puts economists at a disadvantage when talking to legal experts and makes it difficult to build and maintain the internal logic of an argument within a dialogue, in particular if they have to do so *ad hoc*.

In contrast, legal experts achieve precision⁴ in the language they use mostly with the use of legal terms on which there is broadly a common understanding amongst lawyers, with reference to definitions found in the legal text, and by giving meaning to terms through reference to precedent. This is a much more fluid type of precision, as it relies on interpretation rather than mathematics, but just like economics it, too, is contextual.

⁴ Exceptionally, for strategic reasons, the intention may sometimes be the exact opposite: ambiguity. We briefly discuss this below in Section 2(b).

Admittedly, the lawyer wishing to communicate with an economist assisting in a dispute will need to construct arguments using verbal communication without too much reference to technical legal concepts and case law that will be unfamiliar to the interlocutor, just like the economist will have to communicate without reference to mathematical formulae unfamiliar to the lawyer. Both sides would therefore ideally have to make an effort to communicate effectively with each other. Yet, economists face the challenge that panellists and the WTO Secretariat team assisting them are predominantly jurists.⁵ Being outnumbered, it is typically the economist that has to strive to make herself understood in order to convince the lawyers of the validity of their argument rather than *vice versa*.

b) Undetermined variables in equations and constructive ambiguity

The use of formulae or other mathematical representations (e.g. graphs) has the advantage of making very explicit the relationships between different variables (e.g. the benefit of a policy measure in terms of achieving a legitimate policy objective versus the negative effect of the same measure on trade). The tools used by economists are therefore potentially very powerful and can provide very valuable guidance in trade disputes. The precision and transparency that mathematical tools offer, however leads to two disadvantages when it comes to using those tools in the context of a dispute:

- (i) Formulae explain very explicitly what is known but also make very visible what is unknown. In any quantitative work, there will be parameters that need to be estimated, or on which assumptions need to be made. This tends to make legal practitioners hesitant to use the formulae and the economic evidence submitted in such form.⁶ Even though in a legal argument the corresponding type of information may be missing, that fact will be less explicit. Unknowns can be concealed within legal arguments, while they are very visible in a formula. Formulae can also conceal certain parameters and disregard certain factors, as economists necessarily take into account only a limited set of variables. However, in contrast with legal argumentation, what is disregarded in a formula is explicit. Lawyers somehow seem to assume that economics ought to be objective and devoid of any room for interpretation. Misunderstanding what the science of economics is about, they react when they hear from economists that economic evidence, such as studies, rests on certain assumptions or on a certain chosen methodology, forgetting that legal argumentation too functions in a similar way without raising the same doubts as to its value.⁷

⁵ On the usual background of panellists, see C. P. Bown, 'The WTO Secretariat and the Role of Economics in Panels and Arbitrations', in C. P. Bown and J. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement* (Cambridge: Cambridge University Press, 2010), 391-433, at p. 405.

⁶ For instance, in paras. 129-131 of its Report in *Korea – Alcoholic Beverages*, the Appellate Body endorsed the panel's 'scepticism about the consequences of placing undue emphasis on quantitative analyses of the competitive relationship between products'.

⁷ When unknowns are explicit in legal argument, instead, lawyers do not necessarily find this problematic. Making known what is unknown in a given dispute – for example establishing the exact meaning to be accorded to a term – may be the whole point of a certain aspect of the adjudication. In fact, it will often be the case that some treaty provision will have to be interpreted by utilising a specific method of interpretation. No lawyer would doubt the value of arguments put forward in support of a certain desired interpretation of a provision so long as the interpretative conventions internal to the system are used. That is precisely the art of legal argumentation. It is a weighing and balancing exercise, undertaken within the parameters of some established principles and facts, resting on certain assumptions, and sometimes making use of presumptions.

(ii) The second disadvantage is a consequence of the first one. Because formulae are so explicit, their presence in a case makes it difficult for adjudicators to use constructive ambiguity. Sometimes the best result given the circumstances of a specific case will be for the adjudicator to leave some questions open for a later dispute, or to be dealt with by a higher instance adjudicator, or by positive law-making organs through legislation or negotiation.⁸ In fact, when negotiating treaties, lawyers apply a similar technique. Given an impasse, they will try to draft abstract provisions that can be acceptable to all parties, and leave the practical details to be fleshed out later on, if it ever becomes necessary to do so.⁹ Inevitably, too rigid and factual results yielded by formulae will hinder the application of such a technique in adjudication, making their use undesirable for the adjudicator.

c) Applicability: trade economists and legal precedent

Models are developed by academic economists and are intended to apply only in very specific circumstances. It follows that the findings of those models only hold under certain assumptions. In their training, economists do not learn to think about economic findings as being something that has to have general applicability.

Legal treaties – like the WTO Agreements – instead, are meant to be generally applicable to the parties that have ratified them. Many provisions are drafted in the abstract and to the extent that the conditions in the text are met, they will be applicable. Take a general anti-discrimination provision such as the most-favoured-nation principle found in Article I: 1 of the General Agreement on Tariffs and Trade (GATT) 1994 as an example. The provision can be equally applied to coffee¹⁰ as it can be applied to sardines¹¹ (both are “products” within the meaning of the Article), irrespective of the different factual circumstances of the trade flows in those products and the different circumstances that gave rise to the disputes.

For that reason, in the specific context of international economic law disputes, adjudicators engage in treaty interpretation mindful that their ruling may have an effect on future rulings. Hence, a finding, in the sense of a legal ruling, has to be applicable to the particular circumstances of a specific case, but it should also be possible for the *logic* of the ruling to be extrapolated to other disputes.

In order to assist adjudicators in such an exercise, economists would need to be able to see beyond the internal coherence of individual trade models that tend to be only applicable in specific circumstances and to search instead for the coherence of the body of trade models they work with and the rules that govern this body of trade models. Although this is probably possible (see subsection 3(b) in what follows), it is not what a typical trade economist has been trained to do. As a consequence, economists may be hesitant to make

⁸ See, for instance the suggestion of the Appellate Body in *EC – Seal Products*, at para. 5.129: ‘If there is a perceived imbalance in the existing rights and obligations under the TBT Agreement and the GATT 1994, the authority rests with [Members] to address that imbalance.’

⁹ On constructive ambiguity in WTO law, see for instance Panel Report, *Mexico – Telecoms*, at para. 7.3: ‘WTO negotiators sometimes praise the political wisdom of resorting to “constructive ambiguity” as a diplomatic means of enabling consensus on WTO rules’, making reference to M. Moore, *A World Without Walls – Freedom, Development, Free Trade and Global Governance*, (Cambridge University Press, 2003), p. 111.

¹⁰ GATT Panel Report, *Spain – Tariff Treatment of Unroasted Coffee*, L/5135, adopted 11 June 1981, BISD 28S/102.

¹¹ GATT Panel Report, *Treatment by Germany of Imports of Sardines*, G/26, adopted 31 October 1952, BISD 1S/53.

general statements that may affect future rulings, or may end up applying model-specific findings to general set-ups.

d) Hermeneutics: economists and the interpretation of legal texts

As mentioned above, for the economist the meaning of terminology is often linked to a mathematical formula or a specific modelling approach. For the legal expert, instead, terminology tends to be defined in the relevant legal text itself, from a related legal text (especially in civil law jurisdictions) or from case law (especially in common law jurisdictions). Although it is common for adjudicators to have recourse to the context¹², *travaux préparatoires*¹³, and object and purpose¹⁴ of legal provisions, the starting point is always the text of the provision that has to be interpreted and applied.

An interpretative approach that tries to stay as faithful as possible to the text being interpreted, by finding its “ordinary meaning”¹⁵, is particularly prominent in international economic law, as it is heavily influenced by the style of interpretation common in public international law.¹⁶ In certain cases, different language versions are compared.¹⁷ Definitions from dictionaries are almost mechanically¹⁸ inserted in panel and Appellate Body reports and faithfully cited in parties’ submissions.¹⁹

The modern economist is not at all used to an interpretative approach as a means of building an argument. The textual interpretative method applied in law takes away a certain degree of freedom economists are used to having when building an argument or when choosing which model to apply to analyse a specific situation. As a consequence, they find themselves in a situation where they can only employ a restricted set of their toolbox and where they have to do so using their weaker hand, words rather than mathematics.

That economics is limited in this way is something that should best be kept in mind when considering problems of communication between economists and lawyers. Lawyers should not have unreasonable expectations when it comes to the hermeneutic power of an economic analysis of international trade law.²⁰ Likewise, economists should not too zealously attempt to use their science to interpret legal provisions: not all lawyers would subscribe to the idea that all law can be reduced to economics, despite the relative success of law and economics in many fields of law.

¹² Appellate Body Report, *China – Auto Parts*, para. 151.

¹³ Appellate Body Report, *EC – Computer Equipment*, para. 86; Appellate Body Report, *China – Publications and Audiovisual Products*, paras. 404-405.

¹⁴ Appellate Body Report, *US – Shrimp*, para. 153; Appellate Body Report, *EC – Chicken Cuts*, paras. 238-240.

¹⁵ In para. 58 of its Report in *US – Softwood Lumber IV*, the Appellate Body noted that ‘[t]he meaning of a treaty provision, properly construed, is rooted in the ordinary meaning of the terms used’.

¹⁶ Most importantly, Articles 31-33 of the Vienna Convention on the Law of the Treaties. See, in general, Appellate Body Report, *US – Gasoline*, p. 17.

¹⁷ Appellate Body Report, *US – Softwood Lumber IV*, para. 59 and fn. 50.

¹⁸ Appellate Body Report, *US – Gambling*, para. 166.

¹⁹ In fact, an esoteric running joke among lawyers at the Legal Affairs Division of the WTO is how certain dictionaries have almost been elevated to the status of WTO covered agreements! For a recent example of the prolific use of dictionaries see for instance the United States’ Second Written Submission in *US – COOL (Article 21.5)*, available at http://www.ustr.gov/sites/default/files/DS384.US_Sub2_Fin_Public.pdf (last accessed 19 September 2014), containing references to no less than five dictionaries (Exhibits US-22; US-23; US-24; US-25; and US-34).

²⁰ See, for instance, A. van Aaken, ‘Opportunities for and Limits to an Economic Analysis of International Law’, (2011) 3(1) *Transnational Corporations Review* 27.

3. Applying 'Economics' in Disputes

In WTO trade disputes, economic experts are typically pulled into disputes when there is need to work with quantitative methods.²¹ It is by now common practice to use quantitative economic methods in arbitrations. Panels, instead, have hitherto repeatedly dismissed the possibility of relying solely on the quantitative evidence presented by parties.²² This approach does not only stand in a certain contrast to the approach taken by arbitrators. It stands in an even starker contrast to the approach taken by trade policy makers, who tend to be keen on using quantitative economic evidence to back up their political arguments.

In the following we discuss the role two types of quantitative evidence, namely econometrics and simulations, have played in WTO case law. The discussion on the role of econometrics illustrates that panellists have repeatedly used the discussion around assumptions made to generate evidence as an argument to dismiss the evidence that has been presented. In the discussion on the use of simulations, we illustrate the different weight policy makers, arbitrators and panellists appear to give to simulation results when backing up their own arguments.

In trade disputes, econometric techniques have, for instance, been applied to examine whether changes in the price of one product systematically affect the demand of another product. This question is relevant in the context of 'likeness' determinations for a number of non-discrimination provisions in the WTO covered agreements. Econometric analyses are ex-post analyses, whereby historical data are used to identify statistically significant relationships among variables. Simulation techniques, instead, are used for ex-ante evaluations or for assessments of hypothetical situations. Simulation techniques can, for

²¹ These are usually staff members of the WTO Secretariat, even though Article 13.2 of the Understanding on Rules and Procedures Governing the Settlement of Disputes (DSU) gives the power to panels to appoint their own economic experts (see also Annex 4 of the DSU). It has been argued that using external economic experts appointed by the panels could alleviate fallacies in the engagement of panels with complex economic evidence: M. C. Iacovides, 'Marginal Consumers, Marginalized Economics: Whose Tastes and Habits Should the WTO Panels and Appellate Body Consider when Assessing 'Likeness'?', (2014) 48:2 *Journal of World Trade* 323-349, at p. 348. G. Sacerdoti, 'The Nature of WTO Arbitrations on Retaliation', in C. Bown and J. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*, (Cambridge University Press, 2010), pp. 23-33, at p. 32 makes the point that appointment of experts under Article 13 of the DSU should be possible for Article 22.6 of the DSU arbitrations as well. Notably, panels, that undeniably have the possibility of appointing their own economic experts, have hitherto never availed themselves of that possibility: C. A. Melischek, *The Relevant Market in International Economic Law: A Comparative Antitrust and GATT Analysis* (Cambridge University Press, 2013), p. 207. There have been some cases where panels have consulted experts, albeit not economists. These are disputes falling under the SPS Agreement, such as *EC – Hormones, Australia – Salmon, Japan – Agricultural Products II, Japan – Apples, EC – Asbestos*, and *EC – Approval and Marketing of Biotech Products*. For risk assessment in WTO disputes, see in general L. A. Jackson and M. Jansen, 'Risk Assessment in the International Food Safety Policy Arena. Can the Multilateral Institutions Encourage Unbiased Outcomes?', (2010) 35 *Food Policy* 538, pp. 544-546.

²² Appellate Body Report, *Japan – Taxes on Alcoholic Beverages II*, p.25; Appellate Body Report, *Korea – Alcoholic Beverages*, para. 129; Panel Report, *Korea – Alcoholic Beverages*, para. 10.92; Panel Report, *Chile – Alcoholic Beverages*, paras. 7.68-7.69: 'In *Japan – Taxes on Alcoholic Beverages II*, the Appellate Body affirmed the decision of the panel to look at the economic concept of "substitution" as one means of examining relevant markets. However, [it] emphasised that this should be considered together with all other legitimate considerations, in the aggregate [...] the use of cross-price elasticity of demand is not the decisive criterion, but merely one among other criteria [...] [T]he concept of substitution in markets should not be confused or equated with a numerical measurement of the extent of substitution as found in the co-efficient of cross-elasticity [...] [T]he econometric measurement of the degree of substitution may not [...] always adequately reflect the extent of substitution.' (footnotes omitted).

instance, be used to construct a counterfactual and assess how trade flows would have been like in the absence of a certain policy.

a) Econometrics on a bumpy road

Parties have submitted econometric evidence in seven disputes of alleged discrimination under the GATT 1994 and the Agreement on Technical Barriers to Trade (TBT Agreement). These are *Japan – Alcoholic Beverages II*, *Korea – Alcoholic Beverages*, *Chile – Alcoholic Beverages*, *Philippines – Distilled Spirits*, *Thailand – Cigarettes (Philippines)*, *US – COOL*, and *US – COOL (Article 21.5)*. In the first five disputes the evidence submitted related to the issue of substitutability between domestic and imported products, whereas in the two latter disputes it related to actual trade effects of the allegedly discriminatory measure. The evidence submitted was based, in some of the disputes, on rather sophisticated analyses. Yet, in none of the disputes did the submitted evidence play a decisive role for the outcome of the case. In some, the evidence was even dismissed.

In *Japan – Alcoholic Beverages II* and *Chile – Alcoholic Beverages*, time series data were used to produce econometric evidence to show the competitive relationship between the products at stake. In *Japan – Alcoholic Beverages II*, the defendant submitted econometric evidence, in *Chile – Alcoholic Beverages* both parties did. In both disputes, the panels ruled that the products at stake were directly competitive or substitutable, ‘even though the econometric evidence provided could have led to the opposite conclusion’.²³

In *Japan – Alcoholic Beverages II*, the panel noted that the respondent had not managed to rebut the criticisms advanced by the complainants with regard to the presence of auto-correlation and multicollinearity problems in the presented econometric evidence.²⁴ It is well known among economists that analysis of time series data tends to suffer from so-called auto-correlation or multicollinearity problems.²⁵ Notably, the economists that prepared the econometric evidence for the defendant addressed these problems by using no less than 15 different specifications for their regressions. Indeed, it could be argued that auto-correlation and multicollinearity problems were addressed as thoroughly as technically possible in the context of the disputes and with the data available. Nevertheless, the findings of the econometric work were dismissed.

In *Chile – Alcoholic Beverages*, although the econometric evidence showed a low degree of substitution, the panel considered that that level of substitution was sufficient, as what was required was not the quantification of substitution between the domestic and the imported products but simply that there was some substitution, or at least potential substitution.²⁶ Combining the evidence on some (albeit small) substitution with other evidence, the panel then found pisco and foreign distilled spirits to be directly competitive or substitutable. This ruling therefore illustrates that a priori counterintuitive decisions may be taken if no guidelines exist for linking economic evidence to legal terminology.

²³ World Trade Organization, *World Trade Report 2005: Exploring the links between trade, standards and the WTO*, (2005), available at http://www.wto.org/english/res_e/publications_e/wtr05_e.htm (last accessed 24 September 2014), p. 178.

²⁴ Panel Report, *Japan – Alcoholic Beverages II*, para. 6.31.

²⁵ World Trade Organization, *World Trade Report 2005: Exploring the links between trade, standards and the WTO*, (2005), available at http://www.wto.org/english/res_e/publications_e/wtr05_e.htm (last accessed 24 September 2014), pp. 197-198.

²⁶ Panel Report, *Chile – Alcoholic Beverages*, paras. 7.60-7.79.

In *Korea – Alcoholic Beverages*, the complainants submitted historical statistical evidence and the results of a survey based on the responses of consumers following hypothetical changes in the prices of products.²⁷ The panel, while accepting that some of the criticism advanced by the respondent²⁸ was warranted, found that the evidence put forward by the complainants ‘provided useful information regarding at least the potential competitiveness of the imported and domestic products’.²⁹ At the same time, it stressed that ‘quantitative studies of cross-price elasticity are relevant, but not exclusive or even decisive in nature’ since what is at stake is ‘not the degree of competitive overlap, but its nature’.³⁰

In *Thailand – Cigarettes (Philippines)* the complainant submitted three different studies purporting to show that domestic and imported cigarettes were in competition.³¹ Again in this dispute, the panel attached limited probative value to the econometric evidence submitted. It found that the first two studies were inconclusive as they provided evidence compatible with, but not in itself sufficient for establishing ‘likeness’ due to limitations in their design. According to the panel, the switching ratios calculated in the studies could be attributable to factors other than competitiveness between the domestic and imported products.³² When it came to the third study, the panel accepted that it addressed some of the limitations of the other two studies; yet again it found that it was not sufficient in itself to support a conclusion that all domestic products were ‘like’ all imported products.³³

In *Philippines – Taxes on Distilled Spirits*, both parties submitted econometric evidence based on survey data.³⁴ As the data used did not have a time dimension, the econometric exercises conducted did not suffer from the auto-correlation and multicollinearity problems typical in time series analysis. Yet, evidence based on survey data can easily be criticised based on the methodology chosen to conduct the survey and this is what happened in the panel report. The panel concluded that the respondent’s study supported the finding that there was a significant degree of substitution between domestic and foreign products, despite the fact that the authors of the study had concluded the contrary.³⁵ Notably, the panel referred to the complainant’s criticism of the study that ‘the structure of the choices given to respondents creates a starting point bias, as well as some shortcomings associated with the regression model used for the calculations (a mixed logit model)’.³⁶ Generally, the panel followed the approach of previous panels in mentioning the methodological shortcomings of the submitted evidence, while using the evidence to support its findings on ‘likeness’ already determined based on other factors.³⁷

²⁷ Panel Report, *Korea – Alcoholic Beverages*, paras. 5.139-5.142, and 5.143-5.149.

²⁸ See Panel Report, *Korea – Alcoholic Beverages*, paras. 5.210-5.231.

²⁹ Panel Report, *Korea – Alcoholic Beverages*, paras. 10.91 and 10.97.

³⁰ Panel Report, *Korea – Alcoholic Beverages*, para. 10.44.

³¹ Panel Report, *Thailand – Cigarettes (Philippines)*, paras. 7.443-7.446.

³² Panel Report, *Thailand – Cigarettes (Philippines)*, paras. 7.448-7.449.

³³ Panel Report, *Thailand – Cigarettes (Philippines)*, para. 7.450. The conclusions of the panel on ‘likeness’ were not appealed.

³⁴ Panel Report, *Philippines – Taxes on Distilled Spirits*, para. 7.52.

³⁵ Panel Report, *Philippines – Taxes on Distilled Spirits*, paras. 7.112-7.113 and fn. 506.

³⁶ Panel Report, *Philippines – Taxes on Distilled Spirits*, fnn. 432 and 507 (identical wording).

³⁷ See for example, Panel Report, *Philippines – Taxes on Distilled Spirits*, para. 7.76 where the panel stated that ‘[d]espite the shortcomings of the studies submitted by the parties with regard to the competitive relationship between the relevant products, we have also found enough evidence to suggest a significant degree of competitiveness or substitutability for those distilled spirits’. The panel’s findings of likeness were upheld on appeal: Appellate Body Report, *Philippines – Taxes on Distilled Spirits*, paras. 142-157 and 172-173.

The 'likeness' disputes show a general reluctance by the panels to accord considerable weight to the econometric evidence submitted by the parties. A reading of these disputes reveals a similar approach used by the panels. On the one hand, the panels accept that 'likeness' is about the competitive relationship between products and can, thus, be expressed in economic terms. On the other hand, they seem suspicious of the evidence submitted by the parties and unable or unwilling to use it constructively to make a 'scientific' determination of 'likeness' and they emphasise the perceived differences in economists' and lawyers' working methods that we identified in Section 2. By stressing that quantitative evidence of substitution (or lack thereof) is not decisive for the analysis of the competitive relationship between domestic and imported products, the panels retain for themselves extensive discretion to make the determination of the legal question of whether products are 'like' or 'directly competitive or substitutable'. Thus, they preserve the flexibility to manoeuvre skilfully between the criticisms of the econometric evidence advanced by the parties while using the evidence to buttress their conclusions on substitutability.³⁸

US – COOL and the compliance dispute *US – COOL (Article 21.5)* can be distinguished from the rest as the econometric evidence submitted did not relate to 'likeness', which was not an issue in the disputes, but to the actual trade effects of the US labelling measure on imported cattle and hogs. Specifically, the econometric studies that were submitted by the complainants and the respondent purported to show if and how the US measure had affected prices and market shares of imported livestock in the US market.

The panel in these two disputes³⁹ showed a profound understanding of the issues involved. The result is a pedagogic economic analysis of the data provided by the parties⁴⁰; it is clear, reader-friendly, and accessible for the non-specialist, framed in a way that is relevant for the legal issues involved in the case. Interestingly enough, the panel was at pains (in a report that is otherwise exceptionally succinct) to praise econometric approaches as opposed to the shortcomings of descriptive data, as, in the opinion of the panel, the former has the advantage of isolating different factors that may affect the data.⁴¹ What is more, the panel even indicated that a review of economic and econometric evidence and arguments is necessitated by the function of panels as reviewers of fact and law.⁴²

Despite that, the panel reduced the relevance of its thorough discussion of the econometric evidence, by insisting that finding actual trade effects of the measure was not a prerequisite for finding a violation of Article 2.1 of the TBT Agreement.⁴³ Thus, the panel continued with

³⁸ Neven and Trachtman argue that this may be in order for panels and the Appellate Body to retain the power to take into account the non-protectionist purpose (if any) behind the measure while determining 'likeness' since they do not have the possibility of considering it independently: see D. Neven and J. P. Trachtman, *Philippines – Taxes on Distilled Spirits: Like Products and Market Definition*, (2013) 12:2 *World Trade Review* 297-326, p. 306.

³⁹ The panel had the same composition in the original and the compliance disputes.

⁴⁰ The panel's thorough engagement with the econometric evidence may, partly, be explained by the good input of data it had from the parties and from the importance they attributed to econometric evidence: see, for instance, para. 7.443, where the panel notes that 'the parties have *extensively* argued the alleged actual economic and trade effects of the COOL measure ... For this purpose, they have submitted economic figures and analyses as well as econometric studies' (emphasis added).

⁴¹ Panel Reports, *US – COOL*, paras. 7.508-7.509.

⁴² Panel Reports, *US – COOL*, paras. 7.452.

⁴³ Panel Reports, *US – COOL*, paras. 7.442-7.445. On appeal, the Appellate Body endorsed this approach: Appellate Body Reports, *US – COOL*, para. 325. See also, Panel Reports, *US – COOL (Article 21.5)*, para. 7.183. The reasons why the panel chose (or had) to do so are further explored below in Section 3(a)(iii).

the tradition from previous panel and Appellate Body decisions not to accord the econometric evidence a decisive role for reaching its conclusion.⁴⁴ As such, the panel's detailed engagement with the data may be interpreted as a demonstration to the Appellate Body that panels are capable and willing to engage with econometric evidence and 'assess the robustness'⁴⁵ of studies provided by the parties.

Notwithstanding the positive disposition of the panel in *US – COOL* and *US – COOL (Article 21.5)* towards econometric evidence, its approach may be considered problematic. The signal it sends to the parties is that even when a panel is provided with good data and studies by the parties and even when it is comprised by panellists who are able – assisted, no doubt, by the WTO Secretariat team supporting them – and willing to engage constructively with the econometric evidence, nothing will turn on that evidence at the end of the day. The result is that it becomes unclear why parties in future disputes should go through the considerable cost associated with producing econometric evidence and equally unclear why panellists and the WTO Secretariat should devote their very limited time and resources to analysing econometric evidence that is not influential for the outcome of a dispute.

b) Simulations: a relatively welcome guest

Simulation methods are more complex and based on a larger set of assumptions than econometric methods. In order to simulate the counterfactual scenario (i.e. 'what would have happened if...'), economists need to use information of past or standard behaviour of the relevant variables, i.e. the type of information that is typically gathered through an econometric exercise. Using that information, the effect of a policy change is estimated by simulating the alternative path that the relevant variables would take as a result of the policy change. In order to estimate this alternative path, a simulation model may have to make assumptions regarding the way in which economic actors, like producers and consumers, react to the policy change.

If panellists have struggled to base decisions on econometric evidence – as discussed above – it should in principle be expected that they would find it even harder to use information based on simulations that rely on a larger set of assumptions and often use econometric evidence as an input. In trade law, evidence based on simulations has, nevertheless, been used both in disputes and in arbitrations. In trade disputes based on claims of *de facto* discrimination, simulations have been used to determine the effect of a policy measure or the removal thereof on trade, or on price.⁴⁶ In subsidies they have been used to facilitate demonstrating serious prejudice.⁴⁷ In arbitrations, they have helped determine the appropriate level of countermeasures.

⁴⁴ The Panel Report in *US – COOL* was issued only two months after the Panel Report in *US – Clove Cigarettes*. In that case, the panel, interpreting for the first time the meaning of "less favourable treatment" in Article 2.1 of the TBT Agreement had held in para. 7.267 that 'the examination of whether the measure ... provides for "effective equality of competitive opportunities" needs *not* be based on the actual effects of the measure in the market' (emphasis original). The panel in *US – COOL*, would have been aware of the big number of reports that are being appealed: K. Leitner and S. Lester, 'WTO Dispute Settlement 1995-2014 - a Statistical Analysis', (2014) 17 *Journal of International Economic Law*, 191-201.

⁴⁵ Panel Reports, *US – COOL*, paras. 7.513.

⁴⁶ Stochastic simulation models were used in *Australia – Apples* and *US – Shrimp* to assess risk. In the first case, the simulation model was used to calculate the risk of spreading of diseases to Australian apples whereas in the second case to estimate the mortality rate for an Australian loggerhead turtle stock. These two cases fall outside the scope of this chapter and are not discussed in any detail.

⁴⁷ See Articles 5(c) and 6.3 of the SCM Agreement.

Simulations in disputes

In *US – Upland Cotton*, Brazil presented findings from a simulation model to claim that subsidies by the United States had caused serious prejudice to its interests.⁴⁸ Brazil submitted results from simulations performed by an external expert using the Food and Agricultural Policy Research Institute (FAPRI) model that is used in the United States to brief policy makers. The panel decided to take ‘the analyses in question into account where relevant to [its] analysis of the existence and nature of the subsidies in question, and their effects’⁴⁹ but not to rely ‘upon quantitative results of the modelling exercise – in terms of estimating the numerical value for the effects of the US subsidies, nor indirectly in [its] examination of the causal link’.⁵⁰

Brazil submitted the results of a simulation model again in the compliance dispute, with the intention of calculating the price suppression it had suffered as a result of the US subsidy payments. The United States responded by criticising the model, while using the same simulation model by inputting its own values to produce different results. The panel decided to take ‘the analysis of the model one step further’, compared to the panel in the original dispute, ‘by considering in some detail the arguments made by the parties about the model, its assumptions and results.’⁵¹ The panel found that the US measures had led to an increase in US production of cotton that suppressed world prices, although it was not in a position to quantify exactly the counterfactual world price of cotton based on the simulation model.⁵² On appeal, the Appellate Body noted that although the panel had ‘appropriately examined the model, the parameters used by each party, and the arguments made by the parties, and noted the different results generated by the simulations conducted by each party, [it] could have gone further in its evaluation and comparative analysis of the economic simulations and the particular parameters used.’⁵³ Nonetheless, the Appellate Body upheld the panel’s findings and did not find it in breach of its duty to conduct an objective assessment of the matter before it.⁵⁴

A simulation model was presented as evidence by the European Union in *US – Large Civil Aircraft (Second Complaint)*, as part of its causation argument in its claim of serious prejudice under Part III of the Agreement on Subsidies and Countervailing Measures (the SCM Agreement). First, the model had the aim of supporting a general argument that Boeing had used the alleged subsidies to lower its prices on large civil aircraft, and second it purported to quantify the extent to which those subsidies had allowed Boeing to lower its prices.⁵⁵ Following extensive discussion of the model, the panel reached the conclusion that it did not support the existence of a causal link between the receipt of the subsidies and lower prices of Boeing aircraft. The panel based its conclusion on being ‘unable to accept the assumptions on which the model [was] based’ since those assumptions were not ‘an appropriate representation of Boeing’s actual commercial behaviour’.⁵⁶

⁴⁸ This is requirement for actionable subsidies falling under Part III of the SCM Agreement.

⁴⁹ Panel Report, *US – Upland Cotton*, para. 7.1209.

⁵⁰ Panel Report, *US – Upland Cotton*, para. 7.1205. The Appellate Body was ‘not prepared to second-guess the Panel’s appreciation and weighing of the evidence before it’ and did ‘not see any error ... in the application of the law to the facts’: Appellate Body Report, *US – Upland Cotton*, para. 448.

⁵¹ Panel Report, *US – Upland Cotton (Article 21.5 – Brazil)*, para. 10.198.

⁵² Panel Report, *US – Upland Cotton (Article 21.5 – Brazil)*, paras. 10.221-10.222.

⁵³ Appellate Body Report, *US – Upland Cotton (Article 21.5 – Brazil)*, para. 358.

⁵⁴ Appellate Body Report, *US – Upland Cotton (Article 21.5 – Brazil)*, paras. 434-435.

⁵⁵ Panel Report, *US – Large Civil Aircraft (Second Complaint)*, Appendix VII.F.2: The Cabral Model, para. 66.

⁵⁶ Panel Report, *US – Large Civil Aircraft (Second Complaint)*, Appendix VII.F.2: The Cabral Model, para. 76. On appeal, the Panel’s treatment of the simulation model was not specifically addressed.

In *US – COOL*, Canada submitted the results of a simulation model, which measured the impact of the disputed US country of origin labelling measure on the willingness of operators along the supply chain to pay for imported Canadian animals.⁵⁷ The panel noted that ‘the simulation results [could not] be viewed as projections of market price outcomes’ and that the economic model on which the simulation was based was a ‘simplification to illustrate the complex economic mechanism underlying the impact of the COOL measure’, something which the panel thought ‘introduce[d] additional uncertainty into the estimated figures.’⁵⁸ Nevertheless, the panel accepted that the simulation model ‘shed some light on the different types of segregation and compliance costs encountered at different stages of the supply chain’ and how market forces would allocate the costs along the supply chain. Therefore, the panel accepted the simulation model as showing that the US measure was likely to cause a decrease in the volume and price of imported livestock.⁵⁹

In *China – Raw Materials*, the panel found that Article XX of the GATT 1994 was not available to China as a defence for a breach of Paragraph 11.3 of its Accession Protocol.⁶⁰ Examining *in arguendo* whether China had demonstrated that the export duties were justified pursuant to Article XX(b) of the GATT 1994⁶¹, it discussed a simulation model of demand and supply submitted by China, quantifying the impact of its export restrictions on some raw materials⁶², with the aim of showing how its disputed measures made a material contribution to its stated objective of protecting the health of the Chinese population.

The panel noted that it was ‘satisfied with the methodology used by China in making its case’ but had ‘a number of concerns with respect to the reliability of the results of the studies’.⁶³ In particular, the panel was concerned with the domestic supply and demand elasticities used for the estimation of the impact of the export restrictions on manganese, magnesium and silicon carbide, since they were assumed to be the same as the elasticities for another material, namely coke.⁶⁴ Not only was the panel not convinced that there were good reasons to make such an assumption, it also had doubts as to the reliability of the estimation of the elasticities of coke. As a result, the panel concluded that the quantitative estimation of the comparative effects of China’s export restriction on those raw materials was ‘highly speculative’ and unreliable.⁶⁵

Finally, in *US – COOL (Article 21.5)* Canada again submitted a simulation model with the purpose of showing that there were less restrictive non-discriminatory measures that could achieve the objective of the US country of origin labelling rules to the same extent, pursuant to Article 2.2 of the TBT Agreement.⁶⁶ Relying on a variant of the same model, Mexico produced and submitted its own study.⁶⁷ The United States argued that the methodology and theoretical economic model of the Canadian simulation and its Mexican variant were

⁵⁷ Panel Reports, *US – COOL*, para. 7.488.

⁵⁸ Panel Reports, *US – COOL*, para. 7.505.

⁵⁹ Panel Reports, *US – COOL*, para. 7.506.

⁶⁰ Panel Report, *China – Raw Materials*, paras. 7.158-7.160.

⁶¹ Panel Report, *China – Raw Materials*, para. 7.230

⁶² Panel Report, *China – Raw Materials*, paras. 7.519 and 7.527.

⁶³ Panel Report, *China – Raw Materials*, para. 7.528.

⁶⁴ Panel Report, *China – Raw Materials*, paras. 7.529.

⁶⁵ Panel Report, *China – Raw Materials*, para. 7.531. See also para. 7.538 for the panel’s overall conclusion on China’s econometric evidence. On appeal, the Appellate Body upheld the panel’s finding that Article XX of the GATT 1994 was not available to China for a breach of Paragraph 11.3 of its Accession Protocol: Appellate Body Report, *China – Raw Materials*, para. 307.

⁶⁶ Panel Reports, *US – COOL (Article 21.5)*, paras. 7.452-7.454.

⁶⁷ Panel Reports, *US – COOL (Article 21.5)*, para. 7.455.

flawed, inflated, and unrealistic, as they relied on lost export revenue, an approach which, according to it, is unusual in economics and requires sophisticated economic modelling that is sensitive to underlying assumptions.⁶⁸

In this compliance dispute, as in the original *US – COOL* dispute, the panel engaged with the submitted evidence and the parties' arguments in order to assess whether the simulated estimations were reliable.⁶⁹ The panel noted that the different results produced by the Canadian and Mexican models showed how the model was very sensitive to the set of parameters' value considered and agreed with the United States that small changes in the assumptions could have major effects on the simulated figures.⁷⁰ It considered that the complainants had failed to address issues such as the existence of elasticities of substitution between different meats and the use of specific and different supply elasticities for Canada, Mexico, and the United States.⁷¹ Moreover, the panel noted that it was not provided with results based on different assumptions, only on different parameters, meaning that it could not determine and compare findings based on those different assumptions.⁷² For those reasons, the panel decided not to rely on the findings of the simulation models submitted by Canada and Mexico for the assessment of the proposed alternative measures.

Simulations in arbitrations

Article 22 of the DSU provides that when a respondent Member has not implemented the rulings and recommendations of the Dispute Settlement Body (DSB) within a determined reasonable period of time, the complainant Member has the right to retaliate by introducing countermeasures.⁷³ When the respondent Member does not agree as to the proposed level of retaliation, as is most often the case, the matter becomes subject to arbitration. Article 22.6 of the DSU provides arbitrators with the mandate to determine the level of appropriate countermeasures.

Under Article 22.4 of the DSU, there must be 'equivalence' between the countermeasures and the level of nullification or impairment caused by the measure that has been found to be inconsistent with WTO obligations. When it comes to countermeasures in the area of subsidies, according to Articles 4.10 and 4.11 of the SCM Agreement, they need to be 'appropriate', in the sense of proportionate to the nullification or impairment caused⁷⁴, whereas under Article 7.10 of the SCM Agreement, they have to be 'commensurate with the degree and nature of the adverse effect'.

Putting a figure on nullification or impairment presents WTO arbitrators with a problem of great practical importance as they de facto have to generate findings of a quantitative nature. This may explain why to fulfil their mandate of determining whether the level of proposed countermeasures is 'equivalent' (Article 22.4 of the DSU), 'appropriate' (Articles 4.10 and 4.11 of the SCM Agreement), or 'commensurate' (Article 7.10 of the SCM

⁶⁸ Panel Reports, *US – COOL (Article 21.5)*, paras. 7.456-7.458.

⁶⁹ Panel Reports, *US – COOL (Article 21.5)*, para. 7.460.

⁷⁰ Panel Reports, *US – COOL (Article 21.5)*, paras. 7.463-7.464.

⁷¹ Panel Reports, *US – COOL (Article 21.5)*, para. 7.464.

⁷² Panel Reports, *US – COOL (Article 21.5)*, para. 7.465.

⁷³ Alternatively, pursuant to Article 22.2 of the DSU, the respondent Member may, voluntarily, offer compensation. For various reasons, this remains a rarity: B. McGivern, 'Implementation of Panel and Appellate Body Rulings: an Overview', in R. Yerxa and B. Wilson (eds.), *Key Issues in WTO Dispute Settlement: the First Ten Years*, (Cambridge University Press, 2005), pp. 103-105.

⁷⁴ Footnotes 9 and 10 to Articles 4.10 and 4.11 of the SCM Agreement.

Agreement), arbitrators have made increasing use of quantitative economic analysis.⁷⁵ Although generally the use of economic models has been praised in arbitrations, it is also common for the limits of economic analysis to be stressed.⁷⁶

In *US – Offset Act (Byrd Amendment) (EC) (Article 22.6 – US)*, two competing models were put forward by the complainants and the respondent to model the trade effects of the US government's contested transfers. The arbitrator chose to reject both models and construct its own model instead since the complainant's model was too disaggregated and the respondents could not be applied reliably due to lack of data.⁷⁷ The arbitrator's adopted model structure resembled the complainants', but with a higher sectoral disaggregation. In view of the parties' disagreement as to the parameter values that would be used as input in the model, the arbitrator used mid-point elasticity estimates, medium level pass-through, and averages over a three-year period.⁷⁸ The arbitrator noted that economic modelling might lead to imprecise results without the process being rendered meaningless as a result. The arbitrator understood its role as obliging it to engage with quantitative evidence and to construct its own model to 'produce, at a minimum, an outcome which is robust in a lowest common denominator sense, but which is nonetheless, [...] a fair measure of the level of nullification or impairment.'⁷⁹

Simulation models were also used in both *US – Upland Cotton (Article 22.6)* arbitrations.⁸⁰ In the first arbitration, based on Brazil's Article 4.11 of the SCM Agreement claims, Brazil put forward simulation models to calculate interest rate subsidies and the additional export sales obtained by US exporters as a result of the discounts. The arbitrator accepted Brazil's methodology, though it adjusted the results by changing some of the parameters imputed in the equations to better reflect the trade distorting impact of the US subsidies.⁸¹ In the second arbitration, based on Brazil's Article 7.10 of the SCM Agreement claims, Brazil reintroduced the demand and supply log-linear displacement model it had used in the original dispute and the compliance proceedings. The model simulated a counterfactual scenario involving the permanent and anticipated removal of US cotton marketing loans and countercyclical payments. The arbitrator accepted the model and used it to calculate the 'commensurate' countermeasures, using its own set of parameters and inputs.⁸²

⁷⁵ See World Trade Organization, 'Quantitative Economics in WTO Dispute Settlement', in *World Trade Report 2005: Exploring the links between trade, standards and the WTO*, (2005), available at http://www.wto.org/english/res_e/publications_e/wtr05_e.htm (last accessed 2 September 2014), pp. 180-189.

⁷⁶ Y. Renouf, 'From Bananas to Byrd: Damage Calculation Coming of Age?' in C. Bown and J. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*, (Cambridge University Press, 2010), pp. 140-142.

⁷⁷ Arbitral Decision, *US – Offset Act (Byrd Amendment) (EC) (Article 22.6 – US)*, paras. 3.114-3-116.

⁷⁸ Arbitral Decision, *US – Offset Act (Byrd Amendment) (EC) (Article 22.6 – US)*, paras. 3.138, 3.144 and 3.146.

⁷⁹ Arbitral Decision, *US – Offset Act (Byrd Amendment) (EC) (Article 22.6 – US)*, paras. 3.125-3.126.

⁸⁰ For a thorough analysis, see Gene M. Grossman and Alan O. Sykes, 'Optimal' Retaliation in the WTO – a Commentary on the Upland Cotton Arbitration', *World Trade Review*, 10/Special Issue 01 (2011), 133-64.

⁸¹ Arbitral Decision, *US – Upland Cotton (Article 22.6 – US I)*, paras. 4.203-4.278.

⁸² Arbitral Decision, *US – Upland Cotton (Article 22.6 – US II)*, paras. 4.120-4.193.

4. Overcoming communication and interpretation challenges

a) Quantitative evidence: when necessary, communication challenges can be overcome

From the above, two patterns emerge. The first is that an attitude of clear unease and hesitation seems to permeate the handling of econometric evidence by panellists, especially in the issue of 'likeness'. This is in contrast with the approach of arbitrators. Anyone familiar with trade policy, will easily discern also that this attitude is in contrast with how trade policy makers use econometrics. The difference of attitude between arbitrators and panellists is remarkable, given that the former may be the exact same persons as the panellists that served in the original and compliance disputes.⁸³ Secondly, arguably, panellists have been more constructive in their approach to evidence based on simulation models in disputes assessing the effect of a trade policy measure on prices/trade flows than to taking into account econometric evidence in likeness determinations. What, then, could explain the discrepancy between the approach of panellists and arbitrators, and the greater willingness of panellists to engage with simulations than econometric evidence?

Regarding the first pattern that emerges from the disputes, the differences in attitude can to a large extent be explained by the fact that the characteristics of quantitative economic evidence explained in the previous section affect the work of panellists, arbitrators, and trade policy makers in different ways. As argued above, quantitative evidence can – if scrutinized closely – be easily criticized or disputed based on the assumptions made to reach the findings. In policy communications the evidence is nevertheless used, as the policy communicator has the freedom to selectively use evidence that makes his or her case. As long as the findings are not too far-fetched and cannot too easily be discredited, the policy maker is happy to use the evidence in support of his or her policy agenda.

Arbitrators do not pursue a personal policy agenda, but under their mandate they have to meet a legal requirement to quantify 'equivalence'.⁸⁴ Being to some extent compelled to use quantitative evidence, they will also be disposed to use quantitative findings as long as they are based on realistic assumptions that cannot too easily be discredited. For panellists the situation is quite different. Unlike arbitrators, panellists are not explicitly obliged by the text of most provisions in the WTO covered agreements to use or come up with quantitative findings, with the exception of certain provisions relating to subsidies, countervailing duties, and anti-dumping. Unlike policy makers, panellists have to take a neutral position and cannot be seen to one-sidedly defend one position or another.

Arbitrators may feel more compelled to use quantitative evidence since their mandate requires them to come up with some quantitative finding.⁸⁵ Although they probably need to feel that the finding has a certain level of robustness, the fact that they have no choice but to use the evidence in order to quantify the level of impairment puts them in a situation where they have to use quantitative evidence, even if that evidence can, in principle, be disputed. In general, they have demonstrated the ability to use economic models constructively, by accepting their logic and validity, while arriving at different results than the parties by

⁸³ Article 22.6 of the DSU.

⁸⁴ Articles 22.6 and 22.7 DSU and Articles 4.10, 4.11 and 7.10 of the Agreement on Subsidies and Countervailing Measures (SCM Agreement).

⁸⁵ Pursuant to Article 22.7 DSU, the arbitrator '*shall* determine whether the level of such suspension is equivalent to the level of nullification or impairment.' (emphasis added).

inputting different parameters in the equations. Indeed, arbitrators have - like panellists - frequently questioned the parameter assumptions proposed by parties.

Unlike arbitrators, panellists are not explicitly obliged by the text of the WTO covered agreements to use or come up with quantitative findings in disputes involving claims of non-discrimination. The Appellate Body has often affirmed this, as what many of the provisions in the WTO covered agreements protect is competitive opportunities, not actual trade flows.⁸⁶ For that reason, as the Appellate Body sees it, it is not necessary to show *actual* effects on the market in order to prove, for example, a violation of the non-discrimination principles.⁸⁷

The different roles of arbitrators and panellists, and considerations of administrability⁸⁸ can also explain why arbitrators and panellists make different use of quantitative evidence. At the panel level, if one of the parties uses quantitative evidence, the other side will typically immediately try to dismiss that evidence by either presenting its own econometric evidence or by discrediting the other party's evidence. Thus, panellists are placed in a situation where they have to decide which side to 'believe'. In legal terms, the panel needs to decide, on the balance of probabilities, which party has made a better case, and especially whether the complainant has made a *prima facie* case of violation (or non-violation) that has not been rebutted by the respondent.⁸⁹ Thus, while arbitrators have a certain level of leeway to adjust the parameters in a proposed simulation, this is not the case for panellists whose role is to evaluate whether the arguments proposed by one party are more convincing than those proposed by the opponent. As the panel noted in *US - COOL* and *US - COOL (Article 21.5)*, its task - as the case law currently stands - is not to produce a unified economic model that represents the 'correct' one, but rather to assess whether the evidence is robust and reliable.⁹⁰

Given that both the quantitative evidence submitted by one side and the criticism of such evidence submitted by the opposing side tend to refer to mathematical or economic terminology that has meaning that is independent of the legal terminology of the case, panellists have frequently shied away from taking a decision on whose evidence is more probative. Engaging in a thorough analysis of conflicting econometric evidence is cumbersome and time-consuming. Thus, for panellists, an easy (and time-effective) way out is to simply dismiss the quantitative evidence and rely purely on legal argumentation. Finding reasons to discard and disregard the evidence, such as methodological problems,

⁸⁶ See for instance, Appellate Body Reports, *EC - Bananas III (Article 21.5 - Ecuador II) / EC - Bananas III (Article 21.5 - US)*, para. 469.

⁸⁷ For example. Appellate Body Report, *US - FSC (Article 21.5 - US)*, para. 215; Appellate Body Report, *Korea - Alcoholic Beverages*, para. 120; Appellate Body Reports, *US - COOL*, paras. 323-324; and Panel Report, *Korea - Various Measures on Beef*, para. 637. A very good illustration of this is Panel Reports, *US - COOL*, paras. 7.438-7.445, where the panel explains that it is not necessary to find the actual trade effects of the US labelling measure, although it proceeds to actually assess them in the end.

⁸⁸ For instance, the increased amount of time it would take to analyse complex econometric evidence, especially if experts were to be appointed.

⁸⁹ See in general Palmeter and Mavroidis, *Dispute Settlement in the World Trade Organization: Practice and Procedure*, at p. 145 (in relation to evidence and the burden of proof) and p. 155 (in relation to the standard of review) and G. Cook, 'Defining the Standard of Proof in WTO Dispute Settlement Proceedings: Jurists' Prudence and Jurisprudence', (2012) 1(2) *Journal of International Trade and Arbitration Law* 50. See also Appellate Body Report, *US - Wool Shirts and Blouses*, p. 16; and Appellate Body Report, *US - Gambling*, para. 141.

⁹⁰ Panel Reports, *US - COOL*, para. 7.539; Panel Reports, *US - COOL (Article 21.5)*, paras. 7.183 and 7.460.

sampling issues etc., is far more easily administered than engaging with the econometric evidence.

Additionally, panellists do not have any generally accepted benchmark against which to decide whether one econometric finding is more convincing than the other. This absence of a benchmark is particularly striking in the context of likeness disputes. For economists, the assessment of likeness is conceptually very close to the assessment of ‘relevant market’ in competition policy cases. Given that the use of quantitative evidence has become a standard element for the determination of relevant markets in competition policy cases, economists may find it surprising that WTO panels are so hesitant to rely on quantitative evidence for their findings. Yet while several competition authorities have established guidelines on the use of economics in competition cases, such guidelines do not exist for trade law. WTO panellists are thus very much left alone when it comes to deciding which evidence to rely on and which ones to disregard.

This brings us to the second pattern. Why is it the case that panellists seem to have been willing to deal more constructively with simulation evidence – notably in disputes evoking the SCM Agreement – than with econometric evidence in likeness cases? Firstly, the SCM Agreement makes explicit reference to the effects of the subsidy, and panellists are thus expected to decide whether a subsidy has caused “adverse effects”. Like arbitrators, panellists involved in SCM cases may therefore feel compelled to examine quantitative economic evidence. As to other possible reasons, although we can only speculate here, they may be related to the fact that simulations are also used at the arbitration level, a fact that may have made those methods more acceptable in the view of panellists. Perhaps it has also simply been too straightforward in the past to attack evidence provided in likeness disputes on the typical grounds of data quality and methodological problems, something which has created a reputational problem for the use of quantitative evidence in such disputes. This might explain the more positive attitude of the panel in *US – COOL* and *US – COOL (Article 21.5)*, which dealt with actual trade effects rather than likeness. It may also be the case that panellists feel comfortable making determinations of likeness without recourse to econometric evidence and based more on their intuition, whereas they do not feel able to do so when dealing with trade effects. Last but not least, the nature of economic evidence provided and the nature of the provider of the evidence may have played a role in the so-called “big subsidy cases” (*US-Upland Cotton*, *Boeing* and *Airbus*). In all three cases parties made significant investments in modelling and economic analysis and they relied on expert evidence from well-known institutions. This approach stands in a certain discrepancy to the approaches Parties have tended to take in likeness cases.

b) Applying general economic concepts may be easier

It has been argued above that modern economists tend to define concepts through mathematical formulae. Often these concepts are model specific. Yet, there are certain concepts that are rather well accepted among economists and that appear to be sufficiently close to concepts used in trade law, even though the relevant terminology differs across disciplines. In making determinations relating to those concepts, there could be scope for panellists to make more use of economic concepts.

One such concept is welfare maximization.⁹¹ The concept could be very useful when it comes to ‘weighing and balancing’ under Article XX of the GATT 1994.⁹² When maximizing the welfare effect of a policy instrument, economists will compare the different possible marginal effects of such a policy in the so-called first order conditions. If a measure, for instance, has a positive effect on public health, a negative effect on trade and also has a cost, the first order conditions allow economists to assess what a ‘perfect’ policy instrument would look like. This instrument will be the instrument that balances the negative effects of welfare through lower trade and higher public spending (to implement the policy) against the positive effects on public health.

This exercise has strong similarities to the exercise panellists conduct when assessing a national measure under Article XX of the GATT 1994. It will be recalled that such an assessment involves examining the extent to which that measure is necessary to fulfil the objective put forth as justification by the respondent Member (e.g. public health). According to the Appellate Body, ‘determining whether a measure is “necessary” involves a process of weighing and balancing a series of factors that prominently include the contribution made by the measure to secure compliance with the law or regulation at issue, the importance of the common interests or values protected by that law or regulation, and the accompanying impact of the law or regulation on imports or exports.’⁹³ Such a standard incorporates, in essence, a utilitarian calculus in the assessment of the compatibility of Members’ measures with the Article XX of the GATT 1994 exceptions. Seen in this light, necessity (and proportionality) is a legal way to express an economic cost-benefit analysis.

Formulating the arguments as to the necessity of measures in economic welfare terms, parties to the dispute might capture better the interests that are at stake. Think as an example, issues of externalities. Indeed, the Appellate Body has stressed that panels have some latitude in choosing the methodology for the assessment of the necessity of a measure and its contribution to the stated objective and has explicitly mentioned that the methodology may be both quantitative and qualitative.⁹⁴ Although the economic analysis will not be able to provide panellists with a precise answer as to whether a policy is necessary or not, it will provide rather precise indications as to which aspects need to be taken into account in the assessment, which relative order of magnitudes are at play in the assessments, and which implicit assumptions about different effects panellists are making when taking a decision in one or the other direction.⁹⁵ Inspiration as to what values to take into account in this welfare analysis can be drawn from the Preamble of the WTO Agreement.

⁹¹ We recognise, of course, that economic analysis of law based on welfare maximization is not without its critics, both internally (e.g. different strands of law and economics) and externally (e.g. lawyers wanting to rely on values and principles rather than cost-benefit analyses).

⁹² See World Trade Organization, *World Trade Report 2005: Exploring the links between trade, standards and the WTO*, (2005), available at http://www.wto.org/english/res_e/publications_e/wtr05_e.htm (last accessed 24 September 2014), section III.a.

⁹³ Appellate Body Report, *Korea – Various Measures on Beef*, para. 164; Appellate Body Report, *China – Publications and Audiovisual Products*, para. 251.

⁹⁴ Appellate Body Report, *Brazil – Retreaded Tyres*, paras. 145-147 and 151.

⁹⁵ For an example, see C. P. Bown and J. Trachtman, ‘Brazil – Measures Affecting Imports of Retreaded Tyres: A Balancing Act’, (2009) 8(Special Issue 1) *World Trade Review*85.

5. Way forward

In recent years, panellists and arbitrators have frequently engaged in the assessment of economic evidence. In the case of panels, this has notably been the case because of an increased number of Members' submissions containing quantitative economic content, for instance in the form of references to econometric or simulation evidence.

Panel and Appellate Body decisions appear to have struggled to take quantitative evidence submitted by Members on board. We have argued in this paper that this may be due to the absence of a benchmark by which panellists can assess the quality of the quantitative evidence in a situation where both parties provide technical arguments in support or against the evidence that may sound equally convincing or non-convincing to a legal expert (and even sometimes to an economist).

Although panellists have in a number of cases taken a relatively constructive approach towards assessing simulation (as opposed to econometric) evidence, it will often be the case that quantitative economic evidence can relatively easily be criticized because of the parameter values or the methodology chosen to make an assessment. Providing adjudicators with guidelines on how to assess quantitative evidence and benchmarks against which to evaluate the quality of such evidence may be a necessary next step in order to raise the acceptance of the use of quantitative evidence in trade disputes.

In the absence of such guidelines there may, nevertheless, be increased scope for the use of economics in trade disputes. So far, adjudicators appear to have chosen to deal with economic analysis in disputes where quantitative evidence has been submitted by at least one of the parties to the dispute. There is scope to apply economic logic also in cases where no quantitative evidence has been submitted. For instance, when it comes to assessments undertaken under Article XX of the GATT 1994, legal and economic logic appear to be sufficiently close to inspire each other and to jointly raise the quality of an assessment. This is the case, even though, the terminology and the methodology used to express the logic (mathematics versus words) is quite different.

These are the situations in which increased collaboration among legal and economic experts and increased efforts to overcome communication barriers can potentially make a real difference. In this exercise, economists will have to accept that the evidence they provide has to serve the legal arguments and not vice versa. It is therefore for economists to make a convincing case in favour of a stronger and better use of economic evidence in trade disputes. That case will have to be made using words rather than mathematics.