

Japan – Measures Affecting the Importation  
of Apples (AB-2003-4): One Bad Apple?  
(DS245/AB/R): A Comment

DAMIEN J. NEVEN

*Graduate Institute of International Studies, Geneva*

JOSEPH H.H. WEILER

*Jean Monnet Center for International Economic Law &  
Justice, New York University School of Law*

## 1 Introduction

This chapter reviews the decision by the Appellate Body (AB) regarding measures affecting the importation of apples in Japan. Section 2 of the chapter presents some background facts. Section 3 considers the Sanitary and Phytosanitary (SPS) agreement and emphasizes the fact that it imposes a discipline on risk-reducing measures even in the absence of discrimination or protectionism. Section 4 discusses how the evaluation of risk-reducing measures can be undertaken in the context of the SPS agreement. Our discussion focuses on two issues: the scope of the mandate given to the adjudicators and the standard of review that they should apply. We emphasize the difficulty of the task faced by the adjudicators, namely to distinguish between determining the level of risk that a country will find optimal to support (which cannot be challenged) and determining whether risk-reducing measures are necessary to achieve the chosen level of risk. We further observe that the common methodology used by Panels, namely to evaluate the existence of risk in the absence of risk-reducing measures, has limited applicability. We also discuss how this approach can be abused, leading the adjudicators to slip from an evaluation of whether the measures are necessary to achieve a given level of risk to an implicit challenge of the level of risk itself (which should remain the preserve of

the Members). Regarding the standard of review, we argue that a lower standard should be applied to measures that do not threaten fundamental principles like nondiscrimination. Section 5 considers the approach and the findings of the Panel and the AB in light of this discussion. Section 6 discusses the consistency requirement imposed by the SPS agreement regarding risk-reducing measures in different circumstances and argues that it can be a very effective tool to prevent abusive standards, without compromising the autonomy of the States in setting the optimal level of risk that they wish to bear. This section also discusses some of the implications of applying different standards of review to cases that involve discrimination or protectionism and those that do not. Section 7 briefly considers how the Panel and the AB handled methods of risk assessment, and highlights the fact that Japan was held to a very high standard of review. Section 8 discusses the approach of the Panel and the AB toward the precaution principle. We consider the precautionary principle in the context of the SPS agreement and argue that the agreement fits naturally with the distinction between risk and ambiguity and in this perspective allows for one type of rationale behind the precautionary principle (while seemingly excluding others). We also observe that there is at least one issue in which scientific evidence was ambiguous in the case. Accordingly, the Panel's and the AB's unwillingness to apply the precautionary principle in this case can be questioned.

## 2 The dispute's factual matrix

The disease targeted by Japan's phytosanitary measure in this dispute<sup>1</sup> is called "fire blight," often referred to by the scientific name for its bacterium, *Erwinia amylovora* or *E. amylovora*. The AB in its decision offers a useful summation of the case, which we employ here. Fruits infected by fire blight exude bacterial ooze, which is transmitted primarily through wind and/or rain and by insects or birds to open flowers on the same or new host plants. The bacteria *E. amylovora* multiply externally on the stigmas of these open flowers and enter the plant through various openings. In addition to the apple fruit, hosts of fire blight include pears, quince, and loquats, as well as several garden plants. Scientific evidence establishes, as the Panel found, that the

<sup>1</sup> We are relying on the synoptic presentation of the facts by the AB itself in WT/DS245/AB/R of 26 November 2003.

risk of introduction and spread of fire blight varies considerably according to the host plant.

The uncontested history of fire blight reveals significant transoceanic dissemination in the 200-plus years since its discovery. *E. amylovora*, first reported in the New York State in the United States in 1793, is believed to be native to North America. By the early 1900s, fire blight had been reported in Canada from Ontario to British Columbia, in Northern Mexico, and in the United States from the East Coast to California and the Pacific Northwest. Fire blight was reported in New Zealand in 1919, in Great Britain in 1957, and in Egypt in 1964. The disease has spread across much of Europe, to varying degrees, depending on the country, and also through the Mediterranean region. In 1997, Australia reported the presence of fire blight, but eradication efforts were successful and no further outbreaks have been reported. With respect to the incidence of fire blight in Japan, the parties disputed before the Panel whether fire blight had ever entered Japan; but the United States assumed, for purposes of this dispute, that Japan was, as it claimed, free of fire blight and fire blight bacteria. It might be added that Japan wished to remain free of the blight. It should also be noted that Japan's claim that the precise pathway of the pathogen across the Pacific was unknown, was neither contested before nor found to be wrong by the Panel or the AB. It is probably correct that the transoceanic pathway is still unknown. Nonetheless, that claim was, apparently, assumed to be irrelevant.

According to the United States, Japan instituted nine measures to preclude contamination:

- (a) The prohibition of imported apples from US states other than apples produced in designated areas in the states of Oregon or Washington;
- (b) the prohibition of imported apples from orchards in which any fire blight is detected on plants or in which host plants of fire blight (other than apple trees) are found, whether or not infected;
- (c) the prohibition of imported apples from any orchard (whether or not it is free of fire blight) should fire blight be detected within a 500-meter buffer zone surrounding such orchard;
- (d) the requirement that export orchards be inspected three times yearly (at blossom, fruitlet, and harvest stages) for the presence of fire blight for purposes of applying the abovementioned prohibitions;

- (e) a post-harvest surface treatment of apples for export with chlorine;
- (f) production requirements, such as chlorine treatment of containers for harvesting and chlorine treatment of the packing facility;
- (g) post-harvest separation of apples for export to Japan from fruits destined to other markets;
- (h) certification by the US plant protection officials that fruits are free of fire blight and have been treated post-harvest with chlorine; and
- (i) confirmation by Japanese officials of the US officials' certification and inspection by Japanese officials of disinfection and packaging facilities.

The United States claimed that the only products exported from the United States to Japan were “Mature, Symptomless” apples that presented no risk of pathogenic transmission, and that therefore, in relation to this product, the measures in question violated the SPS, notably Arts. 2(2) and 5(1).

The Panel consolidated all Japanese measures into one, but accepted that, although officially only mature, symptomless apples were exported, other immature, symptomatic apples had to be presumed to be part of the trade through fraud and error, and that consequently, the Japanese measures would have to be assessed in relation to both groups.

### 3 SPS — objectives and rationale

This case is not considered doctrinally path-breaking or economically problematic. In it, the AB seemed to consolidate and refine, rather than revise, its previous SPS jurisprudence. It is thus a good case to understand the run-of-the-mill SPS physiognomy. Notable in this case is the fact that, yet again, even such a technologically sophisticated country as Japan was unable to comply with the requirement of risk assessment.

The SPS measures existed, of course, before the adoption of the SPS Agreement. Naturally they would be subject to the traditional GATT disciplines notably Arts, III and XX.<sup>2</sup> In some respects, the

<sup>2</sup> Article XI would, of course, also apply to SPS measures; but as is well known, the combination of the grandfathering provisions of the GATT and the Ad Note to Art. III, meant that most SPS type measures would normally be reviewed in the context of Art. III.

interpretative community of the GATT/WTO – governments, adjudicators, lawyers, and economists, etc. – is “hard-wired” to the underlying objective and rationale of Art. III: separating those State measures that are genuinely instituted to protect against risk to human, animals, and plants from those that, by design or otherwise, are there to protect domestic production and cannot be justified in full or in part on legitimate SPS grounds.

At the heart of the SPS Agreement, and one of the principal sources of the difficulties inherent in its interpretation and application, is its apparent addition to the traditional “discrimination” or “protectionism” rationale of GATT disciplines (which SPS still maintains<sup>3</sup>) of a second “unjustified obstacle” rationale that goes beyond Art. XI measures. The principal locus of the new “obstacle” rationale is to be found in Art. 2(2) SPS and those provisions that flow from it, notably Art. 5.

Members shall ensure that any sanitary or phytosanitary measure is applied only to the extent necessary to protect human, animal or plant life or health, is based on scientific principles and is not maintained without sufficient scientific evidence, except as provided for in paragraph 7 or Article 5.

As stated above, the consequence of Art. 2(2) is that Members may have in place or institute an SPS measure that has neither the object nor the effect of favoring domestic products – indeed, there might not even be a domestic product to compete with the import in question – and yet have such a measure challenged and found incompatible with SPS discipline because, for example, it places an undue burden on producers in the exporting State. If a State cannot show that its phytosanitary measures are really necessary for safety and health, as defined by science, those measures simply cannot stand. Further, the State’s established procedures for risk assessment must now themselves comply with a WTO discipline.

In its field of application, SPS represents a dramatic shift, similar (though not identical) to that occurring in the EC in 1974 with the European Court of Justice’s clarification in the *Dassonville* case that the prohibition of measures having an effect equivalent to quantitative restrictions applied to nondiscriminatory measures. The rationale for

<sup>3</sup> Art. 2(3) SPS.

the broad shift in *Dassonville* and the more limited but similar shift in the SPS agreement is the realization that, in the area of regulatory measures designed to protect against risk, there apparently exist many instances of “inappropriate” State measures that unnecessarily create obstacles to trade without serving the rationale of health or safety, and that these measures may exist without protectionist intent or effect. Note that even under the pre-SPS GATT regime, it was quite common to find State measures failing the test of proportionality and least-restrictive-measure, i.e. they were found to be excessive for their declared purpose. But they were always found to be so in the context of a complaint alleging discrimination and protectionism and as part of the proof of such alleged discrimination and protectionism. Under SPS, a complaint may be brought, as in *Apples* or in *Hormones*, without the need to allege or prove any degree of protectionism or discrimination.

One can wonder about the type of circumstances in which this discipline may apply. First, in the absence of a protectionist motive, the States may have an incentive to apply more stringent standards to imported products than to domestic products. This circumstance arises because the cost of meeting the standards raises the producers’ marginal cost, but the cost increase is typically not fully passed on to consumers (except in extreme cases of perfect competition). Profits also fall. A State will not take foreign profits into account when designing health standards, but it will consider domestic profits. Hence, standards applicable to product categories that are imported will be more stringent (more costly at the margin for a given health benefit) than standards applicable to domestic product categories. Standards applicable to product categories that are both imported and produced domestically will fall in between. Overall, States thus have an incentive to impose higher standards when the cost of meeting the standard is partly borne by foreigners. By imposing some consistency in standards across different product categories (Art. 5.5), the SPS agreement will prevent the States from doing this. Second, the impression is given that the States might burden themselves unnecessarily by adopting excessive or “inappropriate” measures because of deficient methodologies of risk assessment, perhaps because of their limited ability to conduct appropriate risk assessment.

In this essay we want to focus on a particular dimension of the SPS Agreement, namely its application in the context where

neither discrimination nor protectionism is the center of the inquiry.<sup>4</sup> Our underlying thesis is that the implication of the shift from a “discrimination” type of inquiry to an “obstacle” type of inquiry has not been fully worked out by Panels and the AB.

#### 4 On the evaluation of risk and risk-reducing measures

Members are entitled to adopt appropriate SPS measures that reduce risk. But when may a Panel determine that an SPS measure is “inappropriate”? In order to clarify the task of the adjudicators, it is useful to consider a situation involving “risk” as one in which several outcomes could arise in the future. To simplify exposition, we can focus on two outcomes: a bad outcome (*B*) and a good outcome (*G*). Let us also assume that the likelihood that each outcome will

<sup>4</sup> The following is the preamble to the SPS, which articulates its principal objectives:

*Reaffirming* that no Member should be prevented from adopting or enforcing measures necessary to protect human, animal or plant life or health, subject to the requirement that these measures are not applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between Members where the same conditions prevail or a disguised restriction on international trade;

*Desiring* to improve the human health, animal health and phytosanitary situation in all Members;

*Noting* that sanitary and phytosanitary measures are often applied on the basis of bilateral agreements or protocols;

*Desiring* the establishment of a multilateral framework of rules and disciplines to guide the development, adoption and enforcement of sanitary and phytosanitary measures in order to minimize their negative effects on trade;

*Recognizing* the important contribution that international standards, guidelines and recommendations can make in this regard;

*Desiring* to further the use of harmonized sanitary and phytosanitary measures between Members, on the basis of international standards, guidelines and recommendations developed by the relevant international organizations, including the Codex Alimentarius Commission, the International Office of Epizootics, and the relevant international and regional organizations operating within the framework of the International Plant Protection Convention, without requiring Members to change their appropriate level of protection of human, animal or plant life or health;

*Recognizing* that developing country Members may encounter special difficulties in complying with the sanitary or phytosanitary measures of importing Members, and as a consequence in access to markets, and also in the formulation and application of sanitary or phytosanitary measures in their own territories, and desiring to assist them in their endeavours in this regard;

*Desiring* therefore to elaborate rules for the application of the provisions of GATT 1994 which relate to the use of sanitary or phytosanitary measures, in particular the provisions of Art. XX(b).

prevail can be assessed in terms of probabilities, so that  $B$  arises with a probability  $p$  and  $G$  arises with a probability  $(1-p)$ . In some circumstances, it may be difficult to formulate these probabilities, for instance, when there are competing theories regarding the development of a disease. We will discuss below (in Section 8) how decisions should be made in those circumstances, which involve some “ambiguity,” and how the SPS agreement treats such ambiguity. For the time being, we assume that probabilities can be estimated.

It is also useful to distinguish between the level of risk that will prevail in the absence of any risk-reducing measures and the level of risk that a State will consider desirable. This can be formulated in terms of probabilities; let  $p$  be the probability that a bad outcome will arise in the absence of any risk-reducing measures. The level of risk that a State will want to enforce (will consider as optimal) can be denoted  $p^*$ .

In this context there arise two questions, which feature prominently in the case. First, what is the mandate of the Panel? Second, what standard of review should the Panels adopt?

#### 4.1 *The mandate of the adjudicator*

What is the mandate of the Panel? Is the Panel meant to question the level of risk that the State wishes to enforce (i.e.  $p^*$ )? Or is it simply meant to question the measure that has been adopted in order to achieve this level of risk? The AB has consistently insisted that Members retain the autonomy to set their own level of acceptable risk, which may differ from one Member to another (see, for instance, *Australian Salmon*<sup>5</sup> and *Hormones*). Article 4(1) is premised on this primordial understanding:

Members shall accept the sanitary or phytosanitary measures of other Members as equivalent, even if these measures differ from their own or from those used by other Members trading in the same product, if the exporting Member objectively demonstrates to the importing Member that its measures achieve the importing Member’s appropriate level of sanitary or phytosanitary protection.

Obviously, Art. 4(1) contemplates a situation where the levels of SPS protection may differ.

<sup>5</sup> See, e.g. AB in *Australian Salmon*, paragraph 125.

Hence, it appears that the mandate of the Panel is limited to the evaluation of whether the disputed measures are necessary in order to achieve the desired level of risk, and, in particular, whether these measures are supported by scientific evidence.

Before discussing how the Panel and the AB approached this mandate, a few remarks on the level of risk absent any measure ( $p$ ), the optimal (desired) level of risk ( $p^*$ ), and risk-reducing measures (which reduce the probability from  $p$  to  $p^*$ ) may be useful.

- (i) It is clear that the level of risk absent any measure ( $p$ ) will differ among various countries. This is so simply because of differences in local conditions, so that some countries will be less exposed to risk than others (for instance, because of different climatic or geographical conditions). Plants and animals belonging to the same family, as well as humans, may also differ in their susceptibility to certain risk factors in different jurisdictions. It is known, for example, that certain ethnicities are more susceptible to particular diseases than are others.

One also expects that the risk-reducing measures necessary to achieve a given level of desired risk ( $p^*$ ) will differ among countries for the same reasons. Indeed, Art. 2(2) does not, in and of itself, demand or necessarily result in the harmonization of phytosanitary measures across countries. For example, measures pertaining to the type of packaging or treatment of foods necessary to ensure those foods' sanitary condition when put into the stream of commerce might be perfectly safe for a cold climate and yet unsafe for a warm climate, thus requiring different phytosanitary regimes. Of course, the evaluation of risk in the absence of measures, as well as the measures that are in place to reach the desired level of risk, could be subject to dispute, without disputing the standard of safety that the State wishes to afford for its population. For instance, what may be contested and will be in dispute is the extent to which the measure a State adopts to ensure that standard of safety is necessary to achieve that standard, given objective conditions affecting the control of the risk in the importing country. It is assumed that both the importing State and the exporting State have a shared understanding of acceptable risk. It is also not disputed, as might be the case, that the exporting country's phytosanitary regime covering the exports is appropriate for the conditions

in the exporting country.<sup>6</sup> What is claimed is that there are objective circumstances that pertain in the importing country but not in the exporting country, which render unsafe in the importing country a product considered safe in the exporting country.

- (ii) One can also expect that the desired (optimal) level of risk that States will choose to support will differ. The optimal level of risk will result from a tradeoff between the benefits of reducing risks and the cost of risk-reducing measures. The willingness to pay for risk-reducing measures will be determined by preferences, in particular the degree of risk aversion, and resources (the budget constraints that governments face). These factors will likely differ from country to country. For instance, one expects that States under stringent budget constraints will be less willing to pay for a reduction of risk. Some societies have a greater awareness of, and sensitivity to, ecological concerns than other societies. One notices these differences in such mundane policies as roadside billboard advertising, waste-disposal regimes, etc. Sometimes one observes such different sensibilities even in matters of health and safety. Policies towards airbags, seatbelts, and smoking differ from society to society, indicating a complex set of different values reflective of and constitutive of societal identity. One can imagine similar differences in phytosanitary regulation areas. Certain medicines are banned in some countries and not in others. Certain medicines require a doctor's prescription and a trained-pharmacist dispensing service in some countries, but are sold over the counter in other countries.

Local conditions will also affect the cost of risk-reducing measures. Overall, one can thus expect some variance among States in the optimal level of risk that they will choose to support.

- (iii) As States can freely choose their optimal level of risk, nothing would seem to prevent a State from selecting, in principle, a level of zero risk (in the language of the AB itself), i.e.  $p^* = 0$ . However, one should not attach too much attention to this possibility, for at least two reasons. First, as mentioned above, the SPS agreement imposes a consistency requirement on measures in different circumstances. Since the implementation of zero risk

<sup>6</sup> There could, of course, be SPS disputes where the importing country challenges the sufficiency of the measure even for the exporting country, or challenges the manner in which the regime in the exporting country is administered.

would entail a very great cost, which would have to be borne in a variety of different circumstances, it is likely to be prohibitively costly. Second, and more importantly, the notion of a “zero risk” level is an abstraction. It is really a limit case that may not matter very much in practice. Whatever the protective measure in place, it is likely that there will always be a strictly positive probability that a bad outcome will prevail. From this perspective, referring to a “de minimis” notion of risk may be better than employing a notion of zero risk. A “de minimis” risk can be seen as a risk that belongs to a small interval close to zero. The upper bound of this interval could be defined as a level of risk so small that it cannot be measurably affected by risk-reducing measures.

- (iv) The task of the adjudicators is delicate; they will be confronted with risk-reducing measures that are challenged. But they need to distinguish between the level of risk that the States have chosen (which cannot be challenged) – which is presumably ensured by the risk-reducing measures under review – and the necessity of the measures to achieve this level of risk.

Facts are often untidy. Consider the following situations. Imagine two different regimes, under which the same medicine might be permitted for over-the-counter sale in one of the countries, but must be dispensed by an authorized pharmacist in the other. This latter measure is being challenged. The adjudicator will need to distinguish between the following hypotheses: first, this may be an instance of differences in preferences that lead to different optimal risks. If so, the measure should be considered SPS-compliant. Second, it may very well be that the States have the same optimal level of risk but that different risk-reducing measures are necessary to achieve that level of risk in the two countries (for instance, because the general public is less well-educated in the latter). In this instance, the measure should also be considered SPS-compliant. Third, it may very well be that the States have the same optimal level of risk and that the risk-reducing measures have the same effectiveness in the two countries. In this instance, the stricter measure is not necessary and should not be considered SPS-compliant.

- (v) To ascertain the SPS compliance of contested phytosanitary measures requires an examination of whether the *measures* themselves are necessary to ensure health and are based on, and maintained with due regard to, scientific principles and evidence. If the

measures themselves are based on scientific evidence and satisfy all other requirements of the SPS discipline, the only question will be whether they were applied correctly to any given import. But SPS measures are almost invariably contested in the context of a specific dispute where, as stated above, specific products found safe and healthful in one jurisdiction are excluded from another on the basis of a competing set of phytosanitary measures. *De facto*, what the Panels typically end up doing, at least in part — and *Apples* is a good example of this — is to examine whether or not, on the basis of the evidence before them, the products in question are “safe.” That is, in this case, whether or not the actual apples exported to Japan posed a risk.

This approach may be appropriate if considered as a sufficient condition to find that a measure is not SPS-compliant. Indeed, if it is found that the level of risk, in the absence of any risk-reducing measure, is “zero” or falls within a “de minimis” interval, then, a fortiori, risk-reducing measures are not necessary. Any such measure can be deemed non-SPS-compliant.<sup>7</sup>

However, this approach can easily descend into a challenge to the level of risk that the State considers optimal. If we are correct in saying that, in applying Art. 2 and its derivatives, the adjudicators end up evaluating the risk-reducing measures’ compatibility with the SPS by trying to assess whether or not the evidence adduced supports the conclusion that the products in question are safe or unsafe, then it can readily be seen how easy it could be to substitute, as the critical test of safety and consequently of SPS compliance, the risk sensibility of the adjudicator or of the expert witness for that of the importing State.

Sometimes this result is due to a subconscious rejection of the full implications of the regulatory autonomy granted to the States to set the level of optimal risk. This unwillingness or subconscious rejection is sustained, in part, by a certain ambivalence in the SPS toward the notion of real, full autonomy: note, for example, how Art. 4 speaks of the

<sup>7</sup> It could also be argued that evaluating whether there is credible scientific evidence to demonstrate a health risk from the apples, the beef, or the salmon in question is the only method of establishing whether in fact and in law the phytosanitary regulations in question are themselves necessary for protection of health, based on scientific principle, and, especially, supported (i.e. applied as in the contested case at hand) by sufficient scientific evidence.

“*appropriate* level of sanitary or phytosanitary protection” (emphasis added). Could this not be read as an invitation for the WTO adjudicator to find that the *level* of protection itself (rather than the measure put in place to ensure that level) is inappropriate?

Moreover, the vocabulary and, at times, the structure of scientific risk assessment is often such that the distinction between the objective prediction of probability and the value placed on such a prediction is confused. The expert witness will be asked, or the scientific evidence will be perused, to determine the degree of risk that the pathogen will find its way into Japanese commerce, and the answer might be that the risk is “negligible.” Strictly speaking, the scientific evidence should restrict itself as much as possible to the quantification of probability of an occurrence. But typically, experts will be invited to assess the “risk” – an evaluation that might involve a combination of both the scientific probability of an occurrence and a political determination of acceptable danger. The lexical expression “negligible risk” might thus constitute no more than a term for a very low probability, which must then be evaluated in the context of the polity that is assessing the risk. But the term might also fold into it that very value judgment: so small, so negligible, that it is not worth bothering about. That combined judgment is always part of a context that affects the value-judgment element. To an expert from the wildflower-rich country of Switzerland, the risk of contamination of a particular species of flora is 1:10,000,000 and hence “negligible,” and not worth bothering about. If that species were, however, the only species of its kind in some other country, or if it were of considerable economic or cultural significance, the very same probability of 1:10,000,000 might not seem negligible at all.

Hence, the adjudicators – Panel and AB – might end up pronouncing on the reasonableness of the standard of protection of the importing State, rather than on the extent to which science supports the conclusion that the products might compromise that standard. Whereas the adjudicator may appropriately question the good faith of the asserted standard, and whereas the SPS itself stipulates certain conditions for consistency of risk within regulatory areas, ultimately the question of determining the actual degree of risk aversion is meant to be left to the State.

This potential slippage between issues is also important to the extent that it affects certain perceptions and positions of both the adjudicators and the parties. As regards the adjudicators, even though

it is the AB's decision at issue, the relationship between AB and Panel in cases such as this is critical, since the findings as to whether the measures in question are based on scientific principles and are supported by sufficient scientific evidence are matters of fact to be established by the Panel, they should be overturned only with difficulty, and they are the basis on which the AB makes its own findings and issues its Report.

#### 4.2 *The standard of review*

The second issue concerns the standard of review that Panels should adopt while examining SPS measures. In *Hormones*, the European Communities drew a distinction between a “*de novo*” approach and the unhappily termed “deference” approach.

In the view of the European Communities, the principal alternative approaches to the problem of formulating the “proper standard of review” so far as Panels are concerned are two-fold. The first is designated as “*de novo* review.” This standard of review would allow a Panel complete freedom to come to a different view than the competent authority of the Member whose act or determination is being reviewed. A Panel would have to “verify whether the determination by the national authority was “correct” both factually and procedurally.” The second is described as “deference.” Under a “deference” standard, a Panel, in the submission of the European Communities, should not seek to redo the investigation conducted by the national authority but instead examine whether the “procedure” required by the relevant WTO rules had been followed.<sup>8</sup>

We do not believe that “deference” is a particularly useful term, nor do we believe that the alternative to a *de novo* review is simply a review of procedural propriety. Far more helpful are the distinctions drawn in many national administrative- and constitutional-law systems, where the review turns to the strictness of the scrutiny to be applied to a measure. Typically, in judicial review of an administrative action, the adjudicator does not seek to put himself or herself in the position of the administrative authority (*de novo* review) but does more than simply review procedural propriety. The adjudicator will apply some test of reasonableness. An adjudicator will overturn a measure only if it was unreasonable (or even egregiously unreasonable) for an administrative authority to decide as it has. By contrast, if the measure

<sup>8</sup> *Hormones*, AB paragraph 111.

in question threatened a fundamental value, such as a basic, protected constitutional right, a much stricter scrutiny would be applied, seeking a compelling reason to adopt the measure in question. The details differ in different jurisdictions, but these kinds of underlying distinctions are commonly found.

In *Hormones*, the AB gave a singularly unhelpful response to the European Community's contention, by stipulating a nebulous standard based on Art. 11 of the DSU, which requires the Panels to base their findings on an "objective assessment of the facts," whatever that may mean. The AB simply avoided the serious issue underlying the EC submission. This response might have served in the context of a predominantly discrimination-oriented regime ex GATT. But it cannot serve in the differentiated WTO, which has, side by side, discrimination- and obstacle-based disciplines. Article 11 is a catchall phrase that would apply to all disputes and thus does not address what in our minds is the real question. The prohibition against discrimination and protectionism is arguably the most fundamental principle underlying WTO trade disciplines. As such, when a State measure is discriminatory and protectionist, one might argue that it should receive very strict scrutiny with a high burden of justification on the State that promulgates such a measure. By contrast, a nondiscriminatory SPS measure, one might argue, is more akin to an administrative act, and as such should be subject to a lower burden of justification. The failure of the AB to deal seriously with this issue in *Hormones* did not make it go away, as will be seen from our analysis of some of the central issues in the case.

Particular concerns also arise when slippage occurs, namely when the adjudicator, by evaluating the level of risk absent any measure, ends up challenging the optimal risk chosen by the State (as discussed above).

Once the inquiry turns to the degree to which there is or is not scientific evidence to impugn the safety of a product (as an indirect method of evaluating SPS compliance of the State measure that would exclude the product as unsafe), it is very easy to substitute one's own judgment for that of the administrative agency of the importing State. "If," the Panel implicitly reasons, "it has not been proven to *our* satisfaction that there is sufficient scientific evidence to establish a real danger, then *ipso jure* the State measure is not SPS-compliant."

The SPS, like the WTO and GATT more generally, is not altogether helpful on the standard of review to be exercised in such cases. Must it allow the State in question a margin of error? Should not the

question be whether a reasonable government, in promulgating a certain sanitary standard, reasonably come to the conclusion that measure X is necessary to ensure sanitary and phytosanitary protection? This standard of review would mean that only unreasonable (rather than wrong) measures would be struck down. By focusing on the product's safety as a means for determining the measure's compliance with the SPS, there is a risk of upsetting the apple cart, and of WTO Panels putting themselves in the business of government, rather than in the business of reviewing governance. The Panels end up applying a test that, in most countries, administrative tribunals in similar situations would consciously seek to avoid.

### 5 Risk assessment in *Apples*

In *Apples*, the Panel appears to have followed the very methodology outlined above. Its inquiry into the compatibility of the phytosanitary measures put in place by Japan became at its core an assessment of the risk posed to Japan by apples imported from the United States. In relation to each alleged risk, the Panel weighed the scientific evidence presented by the parties as interpreted and augmented by the consulted experts. Once the level of risk was established, it was used as a yardstick to measure the compliance of the phytosanitary instruments with the SPS.<sup>9</sup>

Sometimes the Panel's factual conclusions were categorical:

We therefore conclude . . . that there is not sufficient scientific evidence to conclude that mature, symptomless apples would harbour endophytic populations of bacteria.<sup>10</sup>

On other occasions they were more nuanced:

We conclude . . . that, with respect to mature, symptomless apple fruit, the risk that the transmission pathway be completed is “negligible.” Nevertheless, the experts consulted by the Panel, while firmly considering that the transmission by mature apple fruit is unlikely, suggested . . . that apples from severely blighted orchards . . . not be exported.<sup>11</sup>

One notes here the more qualified terms — “negligible” and “unlikely.” One also notes how the “experts,” and through them the Panel,

<sup>9</sup> Sections 8.177 et seq. of Panel Report.    <sup>10</sup> Section 8.128 of Panel Report.

<sup>11</sup> Section 8.152 of Panel Report.

end up engaging in the business of government, rather than review of reasonableness of governance by government.

We therefore conclude that errors of handling or illegal actions are risks that may be, in principle, legitimately considered by Japan. These risks have been acknowledged by the experts, even though they consider them to be “small” or “debatable.”<sup>12</sup>

There would thus be a small risk that apples containing the fire blight might enter the stream of commerce in Japan despite all precautions. However, on the all-important issue of the existence of a transmission vector from such apples to Japanese fruit, the Panel concluded that:

“... the experts considered the completion of the pathway to be unlikely.”<sup>13</sup>

Of course, one should note, “unlikely” does not mean “impossible.” And the dispositive paragraph states as follows:

We therefore conclude ... that it has not been established with sufficient scientific evidence that the last stage of the pathway (i.e., transmission of the fire blight to a host plant) would likely be completed.<sup>14</sup>

Ultimately, Japan’s claim fails on these findings. Since the risks are minimal, the measures put in place by Japan can easily be found to have no rational relationship to the available scientific evidence.<sup>15</sup> In effect, the Panel did not hold that there was no risk; it held that the risk was so small that it did not justify the measures.

One could interpret the Panel’s finding as suggesting that the risk belonged to a “de minimis” interval. In this perspective, the Panel’s approach and its conclusion that the measure cannot be SPS-compliant would be appropriate.

However, one could certainly question whether the Panel established that the risk fell within a “de minimis” interval. In particular, it is odd that the Panel made a finding that apples from orchards with fire blight should be excluded, even though there was no scientific basis for this.

It is also striking that the Panel never wondered about the optimal level of risk that Japan wanted to enforce. If it had indeed established

<sup>12</sup> Section 8.161 of Panel Report.

<sup>13</sup> Section 8.166 of Panel Report.

<sup>14</sup> Section 8.168 of Panel Report.

<sup>15</sup> Section 8.198 of Panel Report.

that the level of risk absent any measure was within a “de minimis” interval, it would not matter. However, it would seem that it is the Panel’s duty always to inform itself of the optimal level of risk that the State has chosen. This seems essential in order to avoid confusing the level of risk that the measure implies with the question of whether the measure was necessary to achieve this chosen level of risk.

Once in the hands of the AB, given the factual nature of the findings, there is not much the AB is willing or able to do. Thus, in Recital 163 of the AB Report we find:

As we see it, the Panel examined the evidence adduced by the parties and considered the opinions of the experts. It concluded as a matter of fact that it is not likely that apple fruit would serve as a pathway for the entry, establishment or spread of fire blight in Japan.<sup>16</sup> The Panel then contrasted the extent of the risk and the nature of the elements composing the measure, and concluded that the measure was “clearly disproportionate to the risk identified on the basis of the scientific evidence available.”<sup>17</sup> For the Panel, such “clear disproportion” implies that a “rational or objective relationship” does not exist between the measure and the relevant scientific evidence, and, therefore, the Panel concluded that the measure is maintained “without sufficient scientific evidence” within the meaning of Article 2.2 of the *SPS Agreement*. We note that the “clear disproportion” to which the Panel refers, relates to the application in this case of the requirement of a “rational or objective relationship between an SPS measure and the scientific evidence.”

Further, what is not discussed in the case is the application of a standard of review and, in particular, whether as a means for testing SPS compliance one should adopt the approach discussed above: first you check whether, procedurally and substantively, the sanitary and phytosanitary rule was based on science and supported by sufficient scientific evidence. You conduct this review in the manner in which, for example, the French Constitutional Council reviews *ex ante* legislation and not in the American “case and controversy” manner. Then, if necessary, you review, in administrative-law fashion, whether a reasonable state authority applying the rules, which have already been found to be in and of themselves SPS-compliant, could reasonably have reached the result in the specific case. Japan argued, artlessly, that insufficient deference was given to its “approach” to scientific evidence,

<sup>16</sup> Panel Report, para. 8.176.    <sup>17</sup> *Ibid.*, para. 8.198.

a contention summarily rejected. Again, implicitly, the AB endorsed the Panel's methodology, which focused on the evaluation of risk, with all the potential slippage that this methodology entails.

What is never addressed in the legal give-and-take is the weight to be given to the experts' assessment of small risk, likely and unlikely risk, and negligible risk; and, more generally, what the upper bound of the "de minimis" interval could be. There seems to be consensus that there is *some* risk. What weight to be given to this amount of risk is a matter that, one would have thought, should be contextualized. In the assessment of Panel, one has the impression that, despite the existence of some risk, a uniform regulatory rule would be applicable to all.

## 6 Consistency and discrimination

If one is to retain the current orthodoxy according to which, even under the SPS, the Members enjoy autonomy to set their own levels of acceptable risk, it is not, strictly speaking, the task of the Panels and the AB to review that national policy decision. So long as the Members adhere to the various disciplines of the SPS, notably that they engage in an appropriate process of risk assessment and management and that their determination of the risk and the measures to combat it are based on scientific evidence, the review should stop there. The Panel can review the measures for proportionality and hold incompatible with the SPS those provisions that are "more trade-restrictive than required to achieve their appropriate level of sanitary or phytosanitary protection."<sup>18</sup> But what the "appropriate level" of protection is, remains, according to the doctrine developed by the AB in *Hormones* and *Australian Salmon*, a matter for each Member.

The Panels and the AB may, however, as part of their inherent jurisdiction, review the good faith of these determinations by a State. This is a very delicate task – it is never easy for an international tribunal of any kind to base its decision and motivate its decision on the ground of bad faith – that usually amounts to a finding that it simply does not believe the State party in question. It would seem to us that two interconnected devices may render this task somewhat easier.

In the first place, Members in SPS cases should be prodded by the Panels to articulate the level of risk against which the measures

<sup>18</sup> SPS Article 5(6).

in question are designed. In fact, it is difficult to understand how a Member can engage in the kind of risk assessment required by Art. 5 SPS without such an articulation. An inability to articulate the level of risk, even if not in precise quantitative terms, may have significant probative value in indicating that the required risk assessment ex Art. 5 was deficient. In addition, it would be the only way for the Panel to do what falls to it, which is to review whether the measures adopted are indeed scientifically grounded and necessary to ensure or mitigate against the risk level so articulated. In this regard, we respectfully disagree with the ruling of the AB in *Hormones*, in which it held that:

[t]o the extent that the Panel purported to require a risk assessment to establish a minimum magnitude of risk, we must note that imposition of such a quantitative requirement finds no basis in the *SPS Agreement*<sup>19</sup>

The AB may not find an explicit basis in the Agreement, but without some indication of the magnitude of risk aimed at, both Panel and AB risk descending into intuitive approximations.

Members might, of course, put the cart before the horse: examine the measures they have in place and construct from them, retroactively, the level of risk that supposedly informed the choice of the measure in question. This is where the second “bad faith” device may be employed. It is permissible to check the autonomous determination of the level of risk against the requirement of consistency in Art. 5(5) SPS.

With the objective of achieving consistency in the application of the concept of appropriate level of sanitary or phytosanitary protection against risks to human life or health, or to animal and plant life or health, each Member shall avoid arbitrary or unjustifiable distinctions in the levels it considers to be appropriate in different situations, if such distinctions result in discrimination or a disguised restriction on international trade. Members shall cooperate in the Committee, in accordance with paragraphs 1, 2 and 3 of Article 12, to develop guidelines to further the practical implementation of this provision. In developing the guidelines, the Committee shall take into account all relevant factors, including the exceptional character of human health risks to which people voluntarily expose themselves.

In *Hormones*, the Panel gave a rather narrow definition to the meaning of “different situations,” construing the terms to apply to risks

<sup>19</sup> *Hormones*, AB 186.

resulting from the “same substance” and creating the “same adverse health effects.”<sup>20</sup> This would be of limited utility in this dispute, since it is the very absence of fire blight in Japan that informed its measure. Article 5.5 could, however, be read to apply to situations in, for example, a similar sector (in this case safety of plant life,) assuming that all other things are more or less equal. If, for example, as in this case, it were found that Japan insisted on a very low tolerance to risk in relation to fire blight, but in other comparable situations was willing to tolerate far higher levels of risk to the health of plant life, one might be justified in drawing the conclusion, absent a convincing justification from Japan, that its measure was a disguised restriction on trade and hence non-SPS-compliant. The Panel specifically stated in *Hormones* that its ruling there should not be seen as either defining or “further limiting” the meaning of the term “different situation” in Art. 5.5. And the AB confirmed a broader approach:

Clearly, comparison of *several* levels of sanitary protection deemed appropriate by a Member is necessary if a Panel’s inquiry under Article 5.5 is to proceed at all. The situations exhibiting differing levels of protection cannot, of course, be compared unless they are comparable, that is, unless they present some common element or elements sufficient to render them comparable. If the situations proposed to be examined are *totally* different from one another, they would not be rationally comparable and the differences in levels of protection cannot be examined for arbitrariness.<sup>21</sup>

It is to be noted, however, that one cannot apply the consistency clause in the manner suggested above in an overly rigid or mechanical manner. One cannot impugn a set level of risk simply and only because one finds that the State is not consistent in related fields. If this were so, State measures affecting imported products would always have to be pitched at the very lowest standard tolerated in the importing State, militating against, for example, a progressive raising of standards. Article 5.5 should not be construed as a mechanism that would force harmonization to the bottom. But the language of the provision — each Member shall avoid arbitrary and unjustifiable distinctions — could be comfortably used to reject abusive standards without compromising the regulatory autonomy of the State under orthodox understanding. If the State could not explain the maintenance

<sup>20</sup> *Hormones*, Panel 8.176.    <sup>21</sup> *Hormones*, AB 217.

of such distinctions, there would be a means of finding them incompatible with SPS without expressly impugning the good faith of the State. We would like to emphasize that the consistency requirement should not necessarily be seen as a device aimed at uncovering crypto-protectionism. Unjustifiable distinctions can eradicate not only the purposively abusive SPS measure but also the unthinking, careless measures that are detrimental to trade without a real social justification. It is important to note that, according to Art. 5.5, arbitrary or unjustified distinctions are impugned not only if they result in discrimination but also if they result in a disguised restriction on trade. In the context of SPS, a disguised restriction on trade need not be in the context of a protectionist scheme.

The principal innovation of SPS was the introduction of a legal discipline that was not rooted in protectionism. A State measure may be found to be SPS-noncompliant even if, as in *Apples*, there is no finding of protectionism and discrimination. As noted by the AB in *Hormones*, a State measure may be found to be GATT-compliant (i.e. not in violation of, for example, Arts. III and XI) and yet still be noncompliant with SPS. At the same time, a measure that is discriminatory or protectionist is, *ipso jure*, also in violation of the SPS. We have argued in this essay that measures which are putatively discriminatory and protectionist merit stricter scrutiny by the adjudicators, and the burden of justification on the State employing them should, necessarily, be higher than in a situation where the disputed measures are not alleged to be discriminatory or protectionist.

If this is so, this would, in the context of SPS, produce a certain legal paradox, the implications of which can only be sorted out as the jurisprudence develops. Imagine two states applying an SPS measure identical to the one we find in *Apples*. In State A, there is very little production of the product in question and, let us stipulate, the economics of the market are such that protectionism is neither the object nor the effect of the measure in question. In State B there is, by contrast, significant production of the product in question, and the effect of the SPS measure in question is to afford protection and raise suspicion as to the aim of the measure. It would seem that the measure in State B would, and perhaps should, receive stricter scrutiny. Assume further that the measure is “truly” compliant. The result could be that, since State A had a lower burden of justification, its SPS measure would be found not to violate the Agreement, whereas State B, with the higher burden of justification, might find its measure

impugned. Hence, the same measure would be SPS-compliant in one country but not in the other. This may be a source of concern in the following sense: State B's measure appears more suspect because of the protective effect. But at the same time, precisely because State B has significant production of the product in question, the consequences of contamination could be much more devastating than in State A, yet State B, with the higher level of risk, ends up with the higher burden of justification and the greater risk of having its measure impugned.<sup>22</sup> This anomaly may be used as an argument against a thesis that would differentiate the burden of justification based on the factors of protectionism and discrimination. But that, in turn, militates against a principle of judicial review, found in most jurisdictions, according to which a public measure that violates a fundamental norm should receive a higher level of scrutiny.

## 7 Risk assessment

Whilst acknowledging that the Japanese risk-assessment exercise studied several possible hosts of fire blight, including the apple fruit, the Panel then found that the risk assessment was not "sufficiently specific" because the conclusion of the assessment did not purport to relate exclusively to the introduction of the disease through apple fruit, but rather more generally, apparently, through any susceptible host/vector. As the AB explained, the Panel also "found the discussion of possible pathways to have 'intertwined' the risk of entry through apple fruit with that of other possible vectors, including vectors considered more likely to be potential sources of contamination than apple fruit," and hence more damaging. Finally, although the assessment noted the *possibility* of entry, establishment, or spread of fire blight through this vector, it did not properly evaluate the *probability* of the occurrence of such events. The result was to invalidate the risk assessment conducted by Japan.

For Japan, the issue was one of methodology in which Members should enjoy discretion provided that the risk could be established. The AB solidly upheld the Panel and insisted, following its ruling in *Hormones*, that the assessment would have to follow the potential

<sup>22</sup> Of course, because of more damaging consequences, country B may wish to select a lower level of optimal risk. Assuming, however, that both countries have selected a "de minimis" risk, the anomaly will remain.

specific pathogens of the disease, and that these would have to be assessed in relation to the contemplated SPS measure in question.

We do not propose to critique in this piece this reasoning of the AB, except to note that the pattern which now emerges from several SPS cases suggests that, absent an international standard that a State might follow, it will be rather difficult for all countries, and notably developing countries, to conduct the kind of risk assessment that would satisfy the stringent methodological requirements stipulated by the AB.

The legal issue concerned here also goes to the standard of review, which likewise is more assumed than discussed. On the one hand, the AB seems to suggest that there is plenty of leeway for alternative methodologies. Thus, in Recital 204 we read:

Contrary to Japan's submission, however, the Panel's reading of *EC — Hormones* does not suggest that there is an obligation to follow any particular methodology for conducting a risk assessment. In other words, even though, in a given context, a risk assessment must consider a specific agent or pathway through which contamination might occur, Members are not precluded from organizing their risk assessments along the lines of the disease or pest at issue, or of the commodity to be imported. Thus, Members are free to consider in their risk analysis multiple agents in relation to one disease, provided that the risk assessment attributes a likelihood of entry, establishment or spread of the disease to each agent specifically. Members are also free to follow the other "methodology" identified by Japan and focus on a particular commodity, subject to the same proviso.

This passage suggests a reasonableness standard rather than a categorical one. But this seems to be negated by the categorical manner in which Japan is to follow the *Australia Salmon* test of 5.1.<sup>23</sup> We do not want to suggest here that the AB was necessarily wrong from a legal point of view. But it does tip the scales considerably against poorer and less

<sup>23</sup> ... a risk assessment within the meaning of Art. 5.1 must:

- (1) *identify* the diseases whose entry, establishment or spread a Member wants to prevent within its territory, as well as the potential biological and economic consequences associated with the entry, establishment or spread of these diseases;
- (2) *evaluate the likelihood* of entry, establishment or spread of these diseases, as well as the associated potential biological and economic consequences; and
- (3) evaluate the likelihood of entry, establishment or spread of these diseases *according to the SPS measures which might be applied*. (original italics) (Recital 121 *Australia Salmon*).

scientifically equipped and sophisticated Members. The powerful Members seem to get at least two, if not three, bites at this apple. First, they will have much greater clout when negotiating international standards, which then will be *de facto* imposed on less powerful Members. Second, it will be difficult for less powerful Members to match the powerful Members' scientific apparatus when their regimes come into conflict. And finally, the powerful Members will, of course, have more resources to attempt to extricate themselves from an uncomfortable international standard.

## 8 The precautionary principle

We have assumed so far that probabilities can be attached to future events. This assumption arises when the mechanisms through which, for instance, a disease can spread are well understood and when the circumstances affecting the spread of the disease can be observed. Hence, the likelihood of each possible outcome can be estimated with confidence. By contrast, a disease may be poorly understood, and there may be competing theories regarding its development, with no evidence to support one or the other. Accordingly, there may be several probability distributions over possible events, depending on which theory is used.<sup>24</sup> As indicated above, there is no scientific explanation behind the spread of the fire blight across the ocean, so this may be such an instance.

Before discussing how the SPS agreement considers such cases and discussing the approach of the Panel, some insight from decision theory on how to proceed under these two sets of circumstances may be useful. In other words, should the uncertainty surrounding the evaluation of probabilities affect decisions?

### 8.1 Risk and ambiguity

The work of Von Neumann and Morgenstern (1944) has shown that when an objective probability distribution can be defined over a set of outcomes, preferences will be linear in probabilities, at least as long as they respect the "independence axiom." That is also to

<sup>24</sup> Taking it for granted that there is no likelihood that can be attached to possible probability distribution. If it were the case, they could be aggregated to yield one distribution.

say that decisions can be formulated as the result of the maximization of expected utility.

Savage (1956) considered a situation where the decision-maker cannot rely on objective probability distribution. He considered a set of possible events and analyzed the choices that an individual would make over alternative gambles that yield different payoffs in these events. He assumed that the individual would always be able to make a choice. He further assumed that the choice between two gambles would not be affected by a modification of the payoffs that accrue when both gambles yield the same payoff (the independence axiom again). For instance, assume that there are two events, H and L, which are not exhaustive, so that Non H and Non L can also arise. The first gamble yields a payoff of 1 if H occurs and 0 if L occurs. The second gamble yields 0 if H occurs and 1 if L occurs. Both gambles give a payoff of “*d*” if neither H nor L occurs. The independence axiom says that a change in “*d*” should not change the choice between the two gambles. Savage showed that from the observation of choices made under these assumptions,<sup>25</sup> one could generate a relationship between events that is nothing but a subjective probability relationship (such that events could be ordered as more or less probable).

This finding has an important consequence: it says that when no underlying probability distribution is available, preferences will be linear in subjective probabilities. Decisions can be formulated as the maximization of subjective expected utility. This also implies that, faced with alternative probability distribution over a set of events, a decision-maker should pick one and behave as if the resulting probabilities were certain.

To illustrate, consider the following experiment, by Ellsberg (1961). There are two urns, each with 100 balls, which can be either red or black. In Urn 1, the proportion is unknown. For Urn 2, it is known that there are 50 red and 50 black balls. An agent is asked to choose between the following bets: bet red, in which case he gets a prize if a red ball is extracted and zero otherwise; or bet black, in which case he gets a prize if a black ball is extracted and zero otherwise. Consider, first, possible bets over Urn 2. Agents will naturally be indifferent as to betting red or black. Asked the same question of Urn 1, agents will typically provide the same answer. This can be seen as a situation

<sup>25</sup> And a couple of additional technical assumptions.

where there are 101 possible theories about the allocation of balls in the Urn. Agents form the subjective assessment that, in the absence of any information to distinguish among the theories, they are all equally likely, and hence attach a subjective probability of 0.5 that a red (or black) ball will be extracted from Urn 1.

In this framework, there is thus nothing special about scientific uncertainty. Precaution is just like protection: optimal risk-reduction efforts can be obtained from a standard cost-benefit analysis (see Gollier, 2001, for a discussion), using subjective probabilities.

Some suspicion about the validity of the framework, however, arises if agents in the example above are asked an additional question, namely whether they prefer to bet red for Urn 1 or bet red for Urn 2. It turns out that most agents prefer to bet red for Urn 2, rather than bet red for Urn 1. This implies that red from Urn 2 is perceived as more likely than red from Urn 1. But if an agent also prefers to bet black for Urn 2 rather than black for Urn 1, non-red from Urn 2 would appear more probable than non-red from Urn 1. This is inconsistent with the notion that the choice of agents reveals probabilities. That is also to say that choices cannot be compared according to their expected utility. As shown by Ellsberg (1961), the problem arises because the axiom of independence is violated.

This experiment indicates that some agents may be ready to pay more for reducing a risk that is more uncertain. Alternative models of decision have also been developed that do not rely on the axiom of independence. For instance, Henry and Henry (2003) describe a model<sup>26</sup> in which agents face different probability distributions over a set of events. Uncertainty is thus described in terms of a family of distributions. In their framework, choices cannot be compared in terms of expected utility (independence is not assumed) but can be compared in terms of a weighted average of the maximum and the minimum of expected utility that obtains across the possible probability distributions. This allows for a representation of preferences in terms of an attitude to risk (the usual risk aversion as depicted by the shape of the utility function) and an attitude toward the uncertainty with respect to the true probability distribution, which is referred to as the degree of aversion toward ambiguity. The latter is represented

<sup>26</sup> Originally developed by Ghirardato et al. (2002).

by the weights that are given respectively to the maximum and the minimum of expected utility.<sup>27</sup>

Finally, it is worth noting that the decisions that agents take in this framework can also be seen as displaying some “precaution.” The aversion of the agents to ambiguity leads them to take some actions in which they give some weight to worst theories.

### 8.2 *Risk, ambiguity, and precaution in the SPS agreement*

As discussed above, the AB confirmed that Japan’s phytosanitary measures were maintained “without sufficient scientific evidence” and hence the measures were not in conformity with Art. 2.2 of the SPS agreement. At the same time, the AB ruled that the phytosanitary measures imposed by Japan were “not imposed in respect of a situation where relevant scientific evidence [was] insufficient,” so that temporary measures could not be justified under Art. 5.7. The AB noted in particular that the Panel had come across an important amount of relevant evidence and that “a large quantity of high quality scientific evidence . . . had been produced over the years and . . . that the experts had expressed increasing confidence in this evidence.” Altogether, the AB thus seems to have considered (i) that there is reliable scientific evidence on the risks involved in the spread of the disease at stake, and (ii) that the evidence confirmed that the risk of having the disease spread through imported apples was small.

Importantly, the AB also reiterated the Panel’s observation that Art. 5.7 “was designed to be invoked in situations where little, or no, reliable evidence was available on the subject matter at issue.”

It is striking that the structure of the SPS agreement, as well as the interpretation of the agreement given by the AB (and Panel), fit with the distinction between risk and ambiguity. It appears in particular that Art. 5.7 can be used when ambiguity is strong.

The “sufficiency of scientific evidence” in Art. 2.2. and the “(in)sufficiency of relevant scientific evidence” in Art. 5.6 thus also appear to refer to different concepts. The use of similar terms could be confusing, and these concepts could be spelled out more clearly. The structure of the test as to whether a measure would be allowed

<sup>27</sup> For instance, the max-min criteria proposed by Gilboa and Schneider (1989) (such that individuals should behave as if the true theory is the one that yields the lowest expected utility) can be shown to display a strong aversion to ambiguity (see Henry, 2002).

could also be clarified: following the terms of the discussion above, a restrictive measure could be allowed if uncertainty could not be confidently characterized in terms of a probability distribution. If uncertainty could be confidently characterized in terms of probability distribution, a restrictive measure might still be allowed if it significantly reduced the occurrence of an event that, according to this probability distribution, is sufficiently likely and sufficiently damaging. Hence, it would appear that any test of whether a measure could be lawful should start with Art. 5.7 (and not Art. 2.2) and question the ambiguity of scientific evidence. A precise evaluation of the measure at stake (under Art. 2.2) would be conducted only if it were concluded that scientific evidence was sufficiently “unambiguous.”

The question of how ambiguity should be measured in practice in cases where consensus cannot be detected is, however, difficult. It would presumably involve a measure of the subset of possible events for which probabilities are (dis-)similar under the range of possible distributions put forward by experts.

The previous discussion indicates that Art. 5.7 of the SPS agreement can be seen as the expression of a precautionary principle. However, it reflects a particular motive for precaution, namely the presence of ambiguity. Yet, as discussed by Gollier (2001), there are other possible justifications for precautionary actions, and those are not explicitly mentioned in this provision.<sup>28</sup>

These other justifications arise in particular from the dynamic nature of scientific uncertainty. When future risks depend on past consumption (as in the case of climate change), the question arises whether preventive efforts should be undertaken today or tomorrow. On the face of it, the expectation that knowledge will improve over time, and hence that actions will be more efficient in the future, would appear to caution against premature actions. This is the argument that is often advanced to justify the US refusal to sign the Kyoto protocol. It is what Gollier refers to as the “learn, then act” principle. However, if knowledge is improved through the observation of the risks themselves, matters may be different. In those circumstances, the observation of damages today should lead to more preventive actions.

<sup>28</sup> Gollier (2001) uses the subjective expected-utility framework. However, the dynamic effects that he investigates would appear to apply in other frameworks, at least at the level of general principles.

*Second*, it may be that postponing preventive actions will also increase future risks; in those circumstances, the prospect of being poorer in the future may lead to a reduction in the amount of pollution today.

*Third*, the extent to which current decisions change flexibility in the future may also be a concern. Some risks may be subject to irreversibilities, and the lack of preventive measures may reduce the options in the future. In other words, there may be an option value in undertaking preventive action, which in principle can be estimated using real option theory.

Hence, the question arises whether the SPS agreement should not reflect those particular circumstances when precautionary actions are particularly appropriate, namely when future risks are increased, flexibility is impaired, and learning proceeds by observation of current trends.

### 8.3 *Precaution in Japan — Apples ?*

As discussed above, the Panel and the AB ruled that the phytosanitary measures imposed by Japan were “not imposed in respect of a situation where relevant scientific evidence is insufficient,” so that temporary measures could not be justified under Art. 5(7). Still, the absence of scientific explanation behind the spread of fire blight across the ocean may be an instance of ambiguity.

*Apples* is thus a good illustration of the rather constricted nature of Art. 5(7) as an expression of a broader and more capacious notion of precaution. For example, there is, on the one hand, considerable scientific evidence on the mechanisms that explain the potential transmission mechanisms of fire blight. On the other hand, there is no accepted, consensus account that actually tracks and explains how fire blight traveled across the oceans to reach Australia or certain Asian countries. A capacious notion of precaution might suggest that, pending the discovery of the actual transoceanic pathway, a State may be entitled to invoke Art. 5(7).

This raises a question of framing. If you frame the question as to the “macro” pathway in its historical context, one would be driven to the conclusion that there was no scientific evidence. If, by contrast, you frame the question as to the specific mechanisms that explain how the pathogen might migrate from a blighted apple to a healthy one, there is evidence. It is not clear why, in this case, the second framework is more appropriate than the first one. Once again, we come back

to the all-important issue of the standard of review. Had the Panel taken the view that its task was not to establish the risk posed by imported apples, but to establish the reasonableness of the Japanese measure and the circumstances of its application to the specific imports from the United States, it would also have asked itself whether it was reasonable for the Japanese government to look, in the context of 5(7), at the first rather than the second framework, as articulated above. Because the Panel slipped into the business of risk assessment itself, rather than assessment of risk assessment, it could come to its conclusion, which the AB uncritically followed.

### References

- Ellsberg, D. (1961). Risk, Ambiguity and the Savage Axioms, *Quarterly Journal of Economics* **75**(5), 643–669.
- Ghirardato, P., Maccheroni, F. and Marinacci, M. (2002). Ambiguity from the differential viewpoint, Working Paper No. 17/2002 ICER.
- Gilboa, I. and Schmeidler, D. (1989). Max-min expected utility with non unique priors, *Journal of Mathematical Economics* **18**, 141–153.
- Gollier, C. (2001). Should We Beware of the Precautionary Principle?, *Economic Policy* **33**, 301–328.
- Henry, M. (2002). Generalized entropy measures of ambiguity and their measurement, mimeo, Columbia University.
- Henry, C. and Henry, M. (2003). Etat de la connaissance scientifique et mobilisation du principe de precaution, *Revue Economique* **54**, 1277–1289.
- Savage, L.J. (1956). *The Foundation of Statistics* Wiley, New York.
- Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.