What If? Tinkering with the Counterfactual: A Comment on US–Washing Machines (Article 22.6-US)*

Edward J. Balistreri[1]* , Petros C. Mavroidis[2], and Thomas J. Prusa[3]

[1]Duane Acklie College of Business Yeutter Institute Chair, University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA, [2]Edwin B. Parker Professor of Law, Columbia Law School, Jerome Greene Hall, Room 734, 435 West 116th Street, New York, NY 10027, USA and [3]Department of Economics, Rutgers University, New Jersey Hall, 75 Hamilton Street, New Brunswick, NJ 08901-1248, USA
*Corresponding author: Email: edward.balistreri@unl.edu

Abstract

Typically, the WTO Arbitrator, when charged with evaluating the permissible level of countermeasures (suspension of concessions), has chosen a counterfactual state of the world where the challenged (illegal) measure has not been adopted at all. The Arbitrator then would calculate the trade lost because of the adopted (illegal) measure, and thus, decide on the level of permissible countermeasures. In *US–Washing Machines* (Article 22.6-US), deviating from this custom, the Arbitrator adopted a different counterfactual, assuming that the complainant had adopted a different, 'reasonable' measure. The Arbitrator then evaluated the trade lost based on the distance between the adopted (illegal) and the 'reasonable' measure and calculated the level of countermeasures. In this paper, we explain the multitude of perils facing dispute settlement if this approach is adopted in future disputes. We also advance a few thoughts on rethinking the workings of the Arbitrator when measuring the level of permissible countermeasures, since similar slippery slopes risk being reproduced in future cases.

1. Facts of the Case

In *US–Washing Machines*,[1] Korea scored a victory when claiming that the US investigating authority had imposed antidumping (AD) duties in a manner not consistent with the WTO AD Agreement. The United States was requested to comply by bringing its measures into compliance with its obligations under the WTO. It did not do so during the reasonable period of time at its disposal.

Subsequently, Korea filed its request for authorization to impose countermeasures (suspend concessions, in WTO legalese).[2] Korea, following a methodology developed in Bown and Ruta

---

*We are grateful to our dear friends, Kyle Bagwell, Chad Philipps Bown, Bill Davey, Bernard M. Hoekman, Henrik Horn, Rodd Izadnia, Damien J. Neven, and Alan O. Sykes, for sharing their wisdom with us, and to the WTO for never ceasing to surprise us (and the world trading community).

[1]*United States – Anti-Dumping and Countervailing Measures on Large Residential Washers from Korea* (*US–Washing Machines*), WT/DS464/AB/R, 7 September 2016.

[2]The term 'suspension of concessions' refers to the habitual mode of retaliation in WTO: the injured party will inflict a damage on the culprit member by raising its tariffs (which, otherwise it cannot do, since it has conceded to the binding of a 'tariff ceiling'). In principle, it does not have to be this way, as Article 22 of the DSU (Dispute Settlement Understanding),

(2010), estimated its damage in the remedy period to be $711million. (para. 1.4 of the report).[3] And if noncompliance continued, Korea proposed the value of countermeasures in future years to be calculated by adjusting that initial remedy value by the annual growth rate of the washing machine market in the United States.[4] The US argued for an alternative approach based on an Armington elasticities model (para. 3.63).[5] The Arbitrator determined countermeasures of $74.40 million.

The Arbitrator came to this determination by using a counterfactual which diverged from both the one that Korea had proposed and also the alternative one that the United States proposed. In Section 2 of this paper, we discuss the report, aiming to shed light on the choice of counterfactual. In Section 3, we explain why, in our view, this report sets a dangerous precedent, whereas in Section 4, we move to a more normative discussion aiming to provide some insurance policy against similar reports in the future. Section 5 briefly recaps.

2. In the Shoes of the Arbitrator

The Arbitrator, typically composed of the original panel members, or, appointed by the Director-General when warranted, will decide on the level of nullification or impairment (N/I) (Article 22.6 of the DSU). In this dispute, the Arbitrator was composed of the original panel members. The original panel had been appointed by the Director-General, since the parties to the dispute could not agree on its composition.[6]

2.1 The Arbitrator's Mandate

The task of the Arbitrator is to come up with a dollar value, which, in its view at least, will observe the legislative guidance embedded in Article 22.4 of the DSU: equivalence between damage inflicted and permissible retaliation. At any rate, it cannot adjudicate a sum higher than what has been requested (even if, in its view, this sum falls short of the damage suffered). Case law, as reported in Bown and Brewster (2017), is clear on this score.

Unlike public international law, WTO case law has consistently held that damages are due from the end of the reasonable period within which the culprit should have complied, even if the illegal act had been committed years earlier. Decisions thus, by the WTO adjudicating bodies, do not have a declaratory but a constitutive effect of illegality.[7]

2.2 In a Bind Because We Live in a World of Recommendations

Article 19 of DSU distinguishes between panel (compulsory) recommendations, and (optional) suggestions. The former do not prejudge the implementing activities to be pursued. The latter

---

which regulates this issue, allows for suspension of concessions other than tariff concessions. In *US-1916 Act (EC) (Article 22.6-EC)*, without explaining too much why, the Arbitrator practically closed the door to suspension of concessions other than tariff concessions. In *EC–Bananas III (Article 22.6-EC)*, the Arbitrator, once again without explaining too much why, made cross-retaliation onerous. Bown and Brewster (2017) provide an excellent survey of all cases involving evaluation of countermeasures, which have consistently taken the form of suspension of tariff concessions.

[3]Decision by the Arbitrator, *United States – Anti-Dumping and Countervailing Measures on Large Residential Washers from Korea – Recourse to Article 22.6 of the DSU by the United States*, WT/DS464/ARB and Add.1, 8 February 2019

[4]Korea made some additional requests, including a request to retaliate against the unlawful imposition of countervailing duties (CVDs). In this paper, we focus on the request with respect to illegally imposed AD duties. In doing that, we will be focusing on the crux of this report, namely, the appropriate counterfactual.

[5]The dollar value of the US' proposed remedy was redacted in the Arbitrator's Report.

[6]*US–Washing Machines*, WT/DS464/5 of 23 June 2014. We cannot know if the Director-General appointed one or more of the panelists, as the WTO documents contain no information to this effect. Any time the Director-General appoints at least one member of the panel, the panel is considered to have been appointed by him/her.

[7]If it had been the other way around, damages would have to be paid from the moment the illegality was committed, as was the case in the GATT years (Mavroidis, 2000).

do. The issuance of recommendation is compulsory, whereas suggestions are optional in two ways: panels do not have to issue them, even when requested to do so; members do not have to follow them, even when issued. Practice shows overwhelming preference for recommendations, and this case was no exception.[8]

The downside is that the Arbitrator must either accept the counterfactual proposed by one of the parties or construct the counterfactual itself. The issuance of a binding suggestion would have relegated this exercise to redundancy.

2.3 Measuring the Level of Retaliation

In what follows, we first discuss the manner in which the level of retaliation has been decided in prior practice, before turning to the decision in *US–Washing Machines (Article 22.6-US)*.

2.3.1 Is a Counterfactual Necessary?

In para. 3.7, the Arbitrator mentions that, in the past, there were N/I proceedings where the level of retaliation had been calculated without recourse to a counterfactual. The Arbitrator, in footnote 42 of the report, lists nine disputes where counterfactuals were used.[9] That list omits three disputes adjudicated under Article 22.6: *US–FSC (Article 22.6-US)*, *US–Offset Act (Byrd Amendment) (Article 22.6-US)*, and *US–1916 Act (Article 22.6-US)*. Yet in these three omitted disputes, a counterfactual was used.

In *US–FSC*, the question was how should the term 'appropriate' countermeasures be understood. The Arbitrator linked it to the level of subsidy paid, as the export subsidy involved therein was, by legislative fiat, illegal. Hence, in the Arbitrator's view, the damage was the amount of subsidy paid, and there was no need to discuss the issue any further.[10] Effectively, the counterfactual was based on the full removal of the subsidy.

In *US–Offset Act (Byrd Amendment)*, the issue concerned future, quantified (and quantifiable) disbursements. The implicit counterfactual was a world where similar disbursements would not occur.

Similarly, in *US-1916 Act* the issue concerned future, quantified (and quantifiable) final judgments and settlements. Consequently, either explicitly or implicitly, in all 12 prior N/I proceedings the Arbitrator has always had recourse to a counterfactual in order to calculate the level of permissible retaliation.

2.3.2 Who Constructs the Counterfactual?

The law is silent on who constructs the counterfactual, as it simply mandates the Arbitrator to observe equivalence between damage and retaliation without any further guidance. In practice, the original complainant will propose a counterfactual. If accepted, then the Arbitrator will use it. For the original defendant to rebut the proposed counterfactual, it must, not only point to alternatives, but further demonstrate that:

> the particular scenario identified by Antigua as the basis for its counterfactual is not such as to accurately reflect the level of nullification or impairment of benefits accruing to Antigua.[11]

---

[8]Mavroidis (2000) discusses the relevant case law with respect to all these issues.

[9]*EC–Bananas III (US) (Article 22.6-EC); EC–Hormones (US) (Article 22.6-EC); Canada–Aircraft Credits and Guarantees (Article 22.6-Canada); US–Offset Act (Byrd Amendment) (EC) (Article 22.6-US); US–Gambling (Article 22.6-US); US–Upland Cotton (Article 22.6-US I); US–Upland Cotton (Article 22.6-US II); US–COOL (Article 22.6-US); US–Tuna II (Mexico) (Article 22.6-US)*.

[10]Case law on this score changed subsequently with *US–Upland Cotton (Article 22.6-US)*, where a trade-effects approach was adopted, and the level of retaliation was linked to the amount of injury suffered. Hence, since *US–Upland Cotton (Article 22.6-US)* recourse to a counterfactual became a necessity.

[11]Ibid., para. 3.23.

If not persuaded by any counterfactual proposed by the parties, then the Arbitrator will construct its own. This is what happened in *US–Washing Machines*.

2.3.3 The Counterfactual Must Be Reasonable

Awards typically include the phrase that 'the counterfactual must be reasonable', or something to this effect.[12] The term 'reasonable' suggests discretion. In para. 3.9, the Arbitrator acknowledges as much, stating nonetheless that it cannot speculate as to what the likeliest counterfactual is. The counterfactual should be a plausible means to comply, which neither over- nor underestimates the level of retaliation.[13] There is, anyway, an upper bound: as already mentioned, it cannot lead to a sum beyond what has been requested (*non ultra petita*), and anyway not beyond and above the sum resulting from the quantification of the damage suffered (para. 3.11):

> Accordingly, a counterfactual that results in suspension in excess of the benefits that are nullified or impaired by the WTO-inconsistent measures would not be plausible or reasonable.[14]

In para. 1.16, the Arbitrator confirms standing case law to the effect that the methodology when constructing the counterfactual is a matter of its own discretion.

2.3.4 Past Practice

The counterfactual in a typical case, when recourse to a counterfactual is made, is a state of the world where the illegal measure does not exist, the illegal measure is removed or revoked. This has been the preferred scenario in all but three cases.

Of immediate interest to us are precisely the three cases where recourse to a 'reasonable' counterfactual was made. In these cases, the counterfactual was a world where another measure (other than the challenged) had been adopted, and not a world, where the illegal measure was simply considered inexistent.

In *US – Section 110(5) Copyright Act*, the EU claim was that the US authorities had not enforced intellectual property (IP) rights, depriving EU right-holders of their dues. The question was what the European Union could expect from the US in terms of enforcement. Strictly speaking, this was not an Arbitral Award issued under Article 22.6 of the DSU. Nevertheless, the function of the Arbitrator (established under Article 25 of the DSU) was akin to that of the Arbitrator operating under Article 22.6 of the DSU: the only question before it was the quantification of the amount of permissible retaliation, as there was no dispute between the parties that the illegality committed by the United States had not been withdrawn. The Arbitrator constructed a state of the world, where enforcement would depend on a cost–benefit analysis: the US authorities would enforce IP rights whenever costs of enforcement did not exceed the sum of royalties paid. They would cut their coat, in other words, according to the cloth (Grossman and Mavroidis, 2007)

TRIPs (Trade-Related Intellectual Property Rights) is the only positive integration agreement in the WTO. Positive action is expected, and the only question was whether the European Union could reasonably expect that the US authorities enforce IP rights, even when the cost of enforcement activities exceeded the revenue from royalties, especially since non-enforcement in similar cases was non-discriminatory. The Arbitrator responded in the negative.

In *EC–Bananas III (Article 22.6-EC)*,[15] the Arbitrator constructed a counterfactual in order to remedy the violation of Article XIII of GATT (allocation of quotas respecting historic shares),

---

[12] Ibid., para. 3.30.

[13] *US–Gambling (Article 22.6-US)*, para.3.27.

[14] This is why the Arbitrator must first calculate the damage suffered. If the requested sum is below this sum, then it cannot decide a level of retaliation above the requested sum. If the requested sum is above and beyond the level of damage suffered, then the Arbitrator cannot decide on a sum above the level of the damage suffered.

[15] *European Communities – Regime for the Importation, Sale and Distribution of Bananas*, WT/DS27;

which was not covered by the waiver in favour of EU imports of bananas of ACP origin.[16] In its view, a tariff quota regime, similar to the one proposed by the defendant but with a lower in-quota tariff was a reasonable counterfactual (paras. 7.1 et seq.). All that the Arbitrator did in this case was to differentiate the quota allocation.

In *US–Gambling (Article 22.6-US)*,[17] the United States had justified its measures banning Internet gambling by invoking public order. The Appellate Body had found that the measures could not be justified, since the US statutes allowed Internet gambling through one measure, the Interstate Horsing Act (IHA). Nevertheless, only domestic suppliers could trade services under IHA. The Arbitrator, for reasons that are hard to understand, constructed a counterfactual, whereby the United States would have to open up IHA to foreign suppliers as well (paras. 3.57 et seq.).

This is a problematic counterfactual. First, as the dissenting opinion correctly pointed out (paras. 3.67 et seq.), what is the guarantee that the opening of one segment of the wider Internet gambling market suffices for the US measure to become legal? In fact, the intuitive response is that the suggested counterfactual is WTO inconsistent: the Arbitrator has artificially divided the Internet gambling market into two sub-markets, imposed openness in a small part of it, accepted total closure in the other, all in the name of avoiding discrimination. Discrimination though, is but one of the two legs to ensure conformity with Article XIV of GATS. The party invoking this provision must also show that the measure is necessary to protect public order. How is it ever necessary that public order, as invoked in this case, justifies the artificial division of the Internet gambling market? Second, one might legitimately question whether, when constructing a counterfactual other than in a world without the illegality present, the Arbitrator has in fact, ipso facto, exceeded its institutional mandate. We explain. All Arbitrators have consistently held that, under Article 22 of the DSU, they are required to provide a number, that is, the maximum extent of permissible retaliation. Their mandate does not extend to findings of (in)-consistency. This finding should have taken place at the compliance-panel stage. When assuming that the illegality has not been committed, they respect their mandate. All they do is calculate the damage by comparing the world with the measure found to be illegal, with a world where the challenged measure had not been adopted. When they construct any other counterfactual, they do not simply provide a number. They provide a number based on the assumption that their chosen counterfactual is WTO-consistent. What is the guarantee for that? And since arbitral awards under Article 22.6 are not appealable, no one can contest the legality of the chosen counterfactual.

With the exception of *US – Section 110(5) Copyright Act*, which is eminently defensible (because, under Article 25 of the DSU, the Arbitrator could, by virtue of the request by the parties to this effect, decide on the consistency of a reasonable counterfactual, and then calculate the amount of permissible retaliation), the other two, and especially the third case, raise more questions than answers. The argument in favour of the approach in the second case (the *EC–Bananas* dispute) is probably its endorsement by the defendant through its prior submissions (albeit at a different quota level). With this in mind, we turn to the discussion of *US–Washing Machines*.

2.3.5 Calculating the 'Reasonable' Counterfactual in US–Washing Machines

US–Washing Machines is the first Article 22.6 dispute involving WTO inconsistent antidumping and countervailing duties. Three key issues emerged:

a) Does WTO consistency require termination of the order?
b) If not, how does the Arbitrator calculate a 'reasonable' alternative WTO consistent duty level?
c) Regardless of the answer to (a) and (b), how does the Arbitrator estimate the counterfactual value of trade in the remedy year (which occurs many years after the duty was initially imposed)?

---

[16]ACP stands for Africa, Caribbean, and Pacific Group of States; in sum, the former colonies of EU member states.
[17]*United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285.

In *US–Washing Machines*, the WTO inconsistent AD measure was the use of zeroing by the US when calculating the AD duties on Korean exporters. In paras. 3.20 et seq., the Arbitrator implies that AD duties would have been imposed even if zeroing had not been used. The Arbitrator goes so far as to state that Korea could not have reasonably expected its benefits under Article 2.4 of the AD to lead to termination of duties (para. 3.21). The Arbitrator thus, thwarted Korea's counterfactual (paras. 3.13 et seq.), which amounted to a state of the world where duties would never have been imposed.

Conversely, the United States had argued in favour of endorsing a reasonable counterfactual, e.g., another US AD measure. The Arbitrator effectively sided with the US proposal: instead of using W-T, it used W-W without zeroing.[18] The US investigating authority, the US Department of Commerce (USDOC), had used W-W during an administrative review of the margins, and then decided to privilege imposition of AD duties using W-T.[19] Korea explicitly objected during the process to the use of W-W as alternative.

Still, the Arbitrator moved to construct a counterfactual using W-W (paras. 3.25 et seq.) and concluded that the AD duty for LG Electronics should be reduced from 13.02% to a [redacted] percent and for Samsung should be should be reduced from 9.29% to 0%. Importantly, these alternative 'WTO consistent' margins are directly reported in the SAS code used by the USDOC to compute the margins; the SAS code reports margins with and without zeroing under the two preferred calculation methods, W-W and T-T, before the USDOC decided to switch from W-W to W-T (para. 3.27). Neither the US nor Korea challenged the validity of the margins reported by the USDOC. Consequently, the Arbitrator did not have to produce the new WTO consistent margins on its own.

2.3.6 Measuring the Value of Lost Imports

We now discuss the difficult task of computing the counterfactual value of trade in the remedy year. The chief challenge for the Arbitrator is that only the actual value of imports in each year is observed. How can one determine what would have happened in the market 'but for' the WTO inconsistent policy?

We begin by discussing the short-run vs. long-run effects of duties. This is an important consideration given that (i) countries often delay bringing a case to the WTO and (ii) WTO dispute proceedings take several years (or longer) once the dispute is initiated. It can easily be the case that the remedy period in an Article 22.6 proceeding is a decade after the duties were initially imposed. In the *US–Washing Machines* dispute the remedy period was five years after the WTO inconsistent duties were originally imposed.

Both the economic models submitted (by Korea and the US) use short-run predictions to make inferences about the long-run impact. While the two models differ, the crucial difference involves the starting point for evaluating the impact. In the case of Korea, the starting point is the subject supplier's share of the US market immediately before the duty was levied. Then, based on work done by Bown and Ruta (2010), Korea adjusts the pre-duty market share for expected price changes (due to the duty). Korea then applies the adjusted market share to the actual size of the US market in the remedy year as its estimate of what exports would have been without the duties. The clear assumption is that the short-run impact on market share is the relevant metric for assessing the longer-run effect.

---

[18]W-W refers to the method where the weighted average of normal value is compared to the weighted average export price. T-T refers the method where the normal value of individual transactions is compared to export price of individual transactions as close in time as possible to the former. Finally, W-T, which was used in this case in the original investigation, refers to the method where the weighted average of normal value is compared to export price of individual transactions, see Mavroidis and Prusa (2018).

[19]It is difficult to say whether the USDOC was influenced by political considerations as well in doing so. The US authorities have been tinkering with zeroing, even though numerous WTO reports have outlawed it, as we recount in Mavroidis and Prusa (2018).

The US proposed a more elaborate Armington elasticities model to assess the impact of the duty; the US model is quite similar to the style of models used by the US International Trade Commission (USITC) in its trade policy analysis for more than two decades. The Armington model allows washing machines from different sources to be imperfect substitutes for one another. The US' Armington model captures not just how the duties affect subject producers but also how domestic consumers and other producers supplying the product respond to the duty and in this respect it captures the broader impacts of the duties. However, perhaps the most critical aspect of the US approach is that the starting point for the US computations is the actual size of the US market in the remedy year, including the actual value of trade for all key suppliers, including subject imports. In effect, the US' position is that the only market information needed to compute N/I is the sales values in the remedy year. According to the US, using the remedy year market shares as the starting point, an Armington elasticities model can be used to predict the increase in subject supplier's sales if the duties were lower.

Both approaches are problematic. Korea assumes that market shares before the duties accurately predict market share in the remedy year. Effectively, Korea is assuming all subsequent market developments, including new entrants or any other advances that might alter the competitive position of the Korea vis-à-vis others, are a result of the duty. For Korea, the immediate short-run impact accurately captures the long-run impact.[20]

The US, on the other hand, argues the market shares in the remedy year are the rightful starting point for evaluating N/I. The problem with the US approach is most clearly seen in the case when (i) the WTO inconsistent duties are so large as to be preclusive (i.e., they drive subject suppliers out of the market) and (ii) the WTO consistent duties are small. In this scenario, the large WTO inconsistent duties will mean that the starting point for evaluating how a reduction in duty will benefit the subject country is a trade value near (or perhaps even) zero. As a result, the US approach will predict that lowering the duties will have a very small impact. This is because in an Armington model (with standard elasticity assumptions) starting from a very low trade value means the gains from compliance are inevitably small.[21]

For example, suppose due to the WTO inconsistent duties subject supply in the remedy year is small (say $100); the US' Armington model will inevitably predict a small N/I value since even if the model were to predict a large percentage increase in subject supply (say, 10,000%) the actual change in trade value will be tiny. In the limit, if subject trade were driven to zero by WTO inconsistent duties, an Armington model will predict the subject suppliers will remain at zero trade even if the duties are lowered.[22]

2.3.7 Using Armington Models to Calculate the Value of Lost Imports

The Armington elasticity model, as formulated by Francois and Hall (1993), is one of the most common techniques for measuring price and volume effects of a change in trade policy on each source country.[23] A typical Armington elasticity model is graphically depicted in for a case where there are three suppliers – domestically produced (denoted with the subscript d), subject imports

---

[20]The Arbitrator also criticized the assumption that Korean and US washers were perfect substitutes.

[21]The Armington model predicts a percentage change in import value, implying the new value is scaled from the starting value. Hence, starting from a trade value of zero will imply no net increase in imports (as zero times any percentage change is zero).

[22]An alternative way to think about it is that the elasticities, which are critical to the performance of the Armington model, are inappropriate at small import values. Most of the empirical evidence used to measure the elasticities relies on aggregated commodities with substantial and persistent import shares. These elasticities might perform reasonably for small price changes around a benchmark with significant import shares, as in the short run just after the WTO inconsistent duties are imposed. Upon arrival at the remedy year, however, these elasticities would be totally inappropriate. They would need to by scaled up (perhaps thousands of fold) in order to show that the elimination of the duties returns us to original market shares. In the case that subject imports are completely driven out of the market, no finite elasticity is sufficient.

[23]The model is referred to as an Armington model based on the demand system, which differentiates products based on region/country of origin. This demand system was originally specified by Armington (1969).
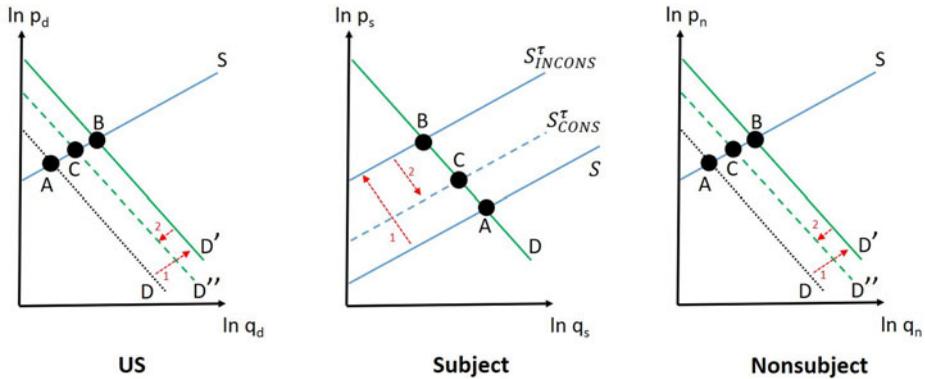
Figure 1. Policy Analysis in an Armington Elasticities Model

(denoted with the subscript s), and non-subject (denoted with the subscript n). For simplicity, one may think of the y-axis measuring price (in USD) and the x-axis measuring the quantity of product. For technical reasons, in Figure 1 we denote the price and quantity in logarithms because that more accurately captures the functional forms used in the actual Armington elasticities model.

The supply and demand curves for each supplier are depicted in their own subfigure. In each chart, the initial equilibrium is given by the point A (where each supplier's demand curve D intersects its supply curve S).

When the United States imposes its WTO-inconsistent duty (denoted as τ^{INCONS}). the relevant supply curve for subject (Korean) exporters to shift from S to S^{τ}_{INCONS} (depicted in the middle chart) and causes the equilibrium to shift from point A to point B (the shift is denoted by the '1' in the subject supplier chart). As a result of the increased duty, subject exporters sell fewer units to the United States and US consumers pay a higher price for those units.

The duty imposed on subject exporters impacts the other suppliers in the market. The demand curve for products sold by (i) US producers and (ii) exporters from the rest of the world both shift out and to the right (depicted as the shift from D to D'); the shift is indicated by the '1' in each chart.

The question for the Arbitrator is, what would be the sales by subject suppliers if the WTO-consistent AD duty were imposed? We denote the WTO-consistent duty as τ^{CONS}, where $\tau^{CONS} < \tau^{INCONS}$).

If the US had adjusted its AD policy so its AD duties were WTO-consistent, the size of the duty would decrease and the relevant supply curve for the subject exporters would shift from S^{τ}_{INCONS} to S^{τ}_{CONS}. This decrease in the duty causes the equilibrium to shift from point B to point C (the shift is denoted by the '2' in the subject supplier chart). As a result, relative to the WTO-inconsistent duty the market outcome (depicted by point B), subject exporters sell more units to the United States and US consumers pay a lower price for those units.

The reduction in the duty on subject exporters also impacts the other suppliers in the market. The demand curve for products sold by US producers and exporters from the rest of the world both shift in and to the left (from D' to D''; this shift is indicated by the '2' in the non-subject and US supplier charts).

While the Koreans submitted a different economic model, the Korean's approach was an attempt to capture the impact on subject sales of moving from point A to point B, and then from point B to point C. Crucially, Korea proposes evaluating the impact using trade values before the duties were imposed and then extrapolates that market impact to the remedy year,

By contrast, the US proposed the Arbitrator use an Armington model to calculate the difference in subject imports between point B and point C using remedy year trade values for the
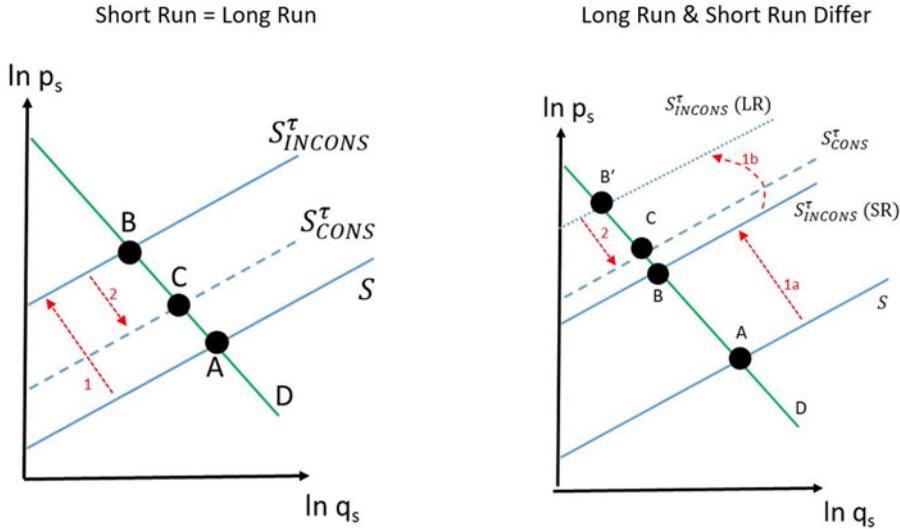
Figure 2. Does Armington Model Capture Actual Market Dynamics?

analysis. The Arbitrator correctly rejected the US approach noting that the trade value in the remedy year reflects the depressing effects of the WTO inconsistent duties imposed many years prior (para. 3.118).

Confirmation that the US' proposed starting point is inconsistent with the underlying Armington model can be seen most clearly if one supposes the WTO-consistent duty were 0% (i.e., a *de minimis* margin). In such a scenario under a properly specified Armington model, the curve S^τ_{INCONS} would shift all the way back to S. In this case, the decrease in duty would cause the equilibrium to shift from point B to point A. In other words, when the tariff is reduced from τ^{INCONS} to 0% the Armington model should predict that subject imports would return to the level they were before the duties were imposed. However, the Armington elasticities model is a short-run model; this prediction only occurs if the starting point for the calculation corresponds to the short-run value.

The conceptual problem with the US' proposed starting point is depicted in Figure 2. Implicit in the US' position is that the model's short-run prediction accurately describes the long-run trade values observed during the remedy period. This is depicted in the left-side panel where the short-run equilibrium (point B) also accurately depicts equilibrium in the remedy year. What occurred in *US–Washing Machines*, on the other hand (and is very likely a concern in any case where many years have passed since the WTO inconsistent duties were imposed) is that market has moved to a new equilibrium, as depicted by point B' in the right-side panel. As depicted, the initial short-run shift in supply is denoted by '1a' and in the longer-run subject supply has experienced an additional shift (denoted by the shift '1b' in the figure). Starting the N/ I counterfactual analysis from point B' would result in a simulated new equilibrium that bears little resemblance to the relevant point implied by the economic theory, point A.[24]

2.3.8 *The Arbitrator's Solution – A Two-Step Armington Model*
While the Arbitrator recognized Korea's view that the market shares in the remedy year were distorted and that consequently that data from the remedy year alone were insufficient to analyse the

---

[24]An alternative interpretation of Figure 2 is that the elasticities used in the Armington model capture short-run effects, not long-run effects. As a result, the model would work if the starting point (in the remedy year) were point B. But, the starting point is point B'.

effects of WTO inconsistent duty, the Arbitrator felt the analytical approach proposed by Korea was too simple. The Arbitrator felt the Armington model proposed by the US was more flexible. Nevertheless, the US model suffered from the serious problem of producing short-run predictions when what was needed was a model that estimated long-run market outcomes. The Arbitrator resolved these issues by creating its own modelling approach – a two-step Armington model.

As the first step, the Arbitrator applied an Armington elasticities model to the US market as it existed *prior to the imposition* of the WTO-inconsistent duties. This first step estimates the impact of imposing the WTO-inconsistent duties on the sales of the subject products by each of the suppliers.[25]

As the second step, the Arbitrator then applied the counterfactual market shares of each of the suppliers (simulated under the first step) to the actual 2017 total value of the US market. This produces a counterfactual estimate of the value supplied by each of the suppliers in 2017 under the assumption that WTO-inconsistent duties are applied. This intermediate calculation is an attempt by the Arbitrator's to isolate the pure effect of the duties and abstract from other changes that have occurred in the US market over the period of time that the WTO-inconsistent duties have been imposed.

With the counterfactual 2017 sales as the starting point the WTO-inconsistent duty is replaced with the WTO-consistent duty and a second Armington policy simulation is performed. This second step simulates the impact of imposing the (lower) WTO-consistent duty on the sales of the supply sources.

The Arbitrator then computes N/I by subtracting Korea's counterfactual 2017 export value with WTO-inconsistent duties (which are the result of step one) from Korea's simulated 2017 export value with WTO-consistent duties (which are the result of step two).[26] Importantly, under the Arbitrator's two-step method, N/I is based on the difference between two counterfactual estimates. Korea's actual Washer sales in 2017 do not enter the N/I calculation.

The Arbitrator argued the two-step approach results make them a more reasonable proxy for long-run results than either Korea's or US's proposals. The crucial innovation is the use of the immediate (or short-run) simulated market shares as a benchmark from which to evaluate the N/I. In addition, if one assumes the key difference over time can be captured by the change in the size of the US market, the results from the Armington elasticities model can be conveniently scaled to match the size of the market in the remedy year.

3. Beware When Tinkering with the Counterfactual

This award is not unproblematic. In fact, any time a counterfactual is chosen, problems emerge. Problems did emerge in past cases as well where a reasonable' counterfactual was chosen. We explain.

3.1 The Irresistible Appeal of Simplicity

The simplest counterfactual is the case where the challenged measure is presumed never to have been adopted. In this scenario, econometric evidence will be typically used in order to construct a hypothetical market that describes how the suppliers (and buyers) would have acted if the challenged measure had never been present. Yes, speculation even in this scenario, is unavoidable. This is nevertheless, the only inconvenience.

---

[25]In *US–Washing Machines*, the Arbitrator only considered two varieties, the subject Korean washers and an 'all others' category that included domestic production and all non-subject imports (this was how Korea proposed dividing the suppliers). However, in general one would expect the Arbitrator to consider at least three varieties: subject imports, non-subject imports, and domestic production.

[26]Korea's export value is calculated at world prices (net of duty). The duty payments do not accrue to Korea and are therefore irrelevant to complainant damages.

For example, consider the growth rates that were used to compute the counterfactual in the *EC–Bananas* dispute. When constructing the ex post counterfactual, say in 2000, for damage suffered in the years 1995–2000, the Arbitrator could have used actual growth rates observed in the bananas market or, alternatively, could have estimated growth rates. In the former case, the presumption is that growth rates would be the same irrespective of whether the challenged measure was ever adopted. In the latter case, the Arbitrator would assume absence of the challenged measure and then use statistical techniques to calculate the hypothetical growth rate. At any rate, some speculation is inevitable. It can, of course, be somewhat tempered, depending on the robustness of the methodology used.

On the other hand, there is an undeniable advantage to computing N/I assuming the policy is fully removed/revoked: there is no doubt about the legality of the counterfactual. One cannot be confident this is true when recourse to a 'reasonable' counterfactual, other than absence of commission of the illegal act, has been made.

The counterfactual used in *US–Gambling* highlights our point. The only way the counterfactual used in *US–Gambling* could be WTO consistent would be if the IHA exhausted the forms of Internet gambling. But we know that this was not (and is not) the case. The United States, thus, continues to violate the WTO (since it imposes restrictions when it had not made any contractual provision to this effect for all forms of Internet gambling), while paying compensation only for a small part of it (IHA).

Or, even if the counterfactual were WTO-consistent, it could be that its plausibility is questionable. The Arbitrator, in the two reports cited above (*EC–Bananas*; *US–Gambling*), tried to insulate itself from similar criticism, when stating that it was under no obligation to identified the likeliest counterfactual. It sufficed that the Arbitrator identified one plausible scenario. However, it is unclear whether the counterfactual's plausibility is measured by reference to the Arbitrator's or the parties' expectations.

Again, consider the *EC–Bananas* dispute. There is no doubt that tariff quotas are legal, and were extensively used in the realm of negotiations on farm trade during the Uruguay round. Indeed, the Arbitrator ensured that its counterfactual would not suffer from illegalities (para. 7.7). The Arbitrator went on and imposed one global tariff quota for both third countries, as well as for non-traditional ACP exports.[27] What is the guarantee that this counterfactual is plausible, or something the trading partners of the European Union could have legitimately anticipated?

What was clear during the Uruguay Round was that the EU delegates were keen to preserve the market share of ACP bananas and the ensuing profits of European distributors (Messerlin, 2001). This is why they signed the Framework Agreement and guaranteed preferential rates to bananas originating in ACP countries, which agreed to forego their rights under the MFN (most-favoured nation) clause. Under the circumstances, the plausibility of a global tariff quota can be seriously questioned.

The borderline case is *US – Section 110(5) Copyright Act*. But in fact, this case should not serve as example at all. The mandate of an Arbitrator under Article 25 is not the same as that of an Arbitrator under Article 22.6. In this dispute, there was partial overlap, and this is why we referred to *US – Section 110(5) Copyright Act*. The Arbitrator in *US – Section 110(5) Copyright Act* though, was entitled to construct a counterfactual, as per the request of the parties.

The crux of the matter is this: the legality of implementing measures is discussed during the compliance procedure. During the compliance review, panels will question and evaluate the legality of a measure that has been adopted, and nothing else. If they find that it suffices to implement the recommendation, then it is the end of the story. If not, then we simply have no clue as to the legality of any other measure that could have been adopted, 'reasonable' or not.[28]

---

[27]These quotas cover quantities in excess of traditional quantities supplied by traditional ACP countries (12 countries, identified in the report), and any quantities supplied by ACP countries which are not traditional suppliers of the EU market.

[28]Practice has evolved, arguably against the letter of law, and has allowed for a second bite at the pie so to speak, through the acceptance of a second compliance panel. The DSU of course, does not allow for anything like this to happen, as there is

There is an additional issue to discuss here. What is the likelihood that Korea might have litigated the case, had the USDOC used W-W? What if Korea was keen in outlawing zeroing when used in a W-T scenario, and nothing else? Since standing AB case law holds that zeroing is WTO-inconsistent in a W-W and in a T-T scenario, Korea might have initiated the dispute to peg the final nail in the zeroing coffin. If true, Korea's damage would be exhausted in the legal fees it paid, which cannot be recuperated, as per standing WTO case law (Mavroidis, 2000). We do not know if this was the case, but the Arbitrator does not know that either.

3.2 Institutional Issues

The use of the 'reasonable' counterfactual raises a few institutional issues as well. We discuss them succinctly in what follows.

3.2.1 Is the Panel Engaging in a De Novo Review?

Standing case law suggests that panels cannot engage in a *de novo* review, that is they must take the factual record as established by the investigating authority. They cannot redo the investigation. But, this is exactly what the Arbitrator has done in *US–Washing Machines*. The USDOC had calculated W-W margins (without zeroing) during the investigation, but had based the duties it actually imposed using W-T (with zeroing). The Arbitrator resurrected the W-W calculation, and used it as benchmark, as its 'reasonable' counterfactual, in order to calculate the level of retaliation.

In this case, data about W-W were readily available. What if this had not been the case though? What if the USDOC had computer code had only reported W-T margins? Would the Arbitrator have redone the calculation from scratch? We will never know, alas.

3.2.2 Why is T-T Not Equally Reasonable?

The Arbitrator makes an assertion, an un-explained affirmation to the effect that W-W is a reasonable counterfactual. The Arbitrator further affirms, citing prior case law, that there is no need to search for the likeliest counterfactual. Thus, through an invented standard of review, the Arbitrator makes life easy for itself. What is the guarantee, though, that a plausible counterfactual observes the quintessential mandate of the Arbitrator–ensuring equivalence between damage and retaliation?

Why, for example, in *US–Washing Machines*, is W-W more reasonable than T-T? More generally, what are the statutory underpinnings of the 'reasonableness' standard? Article 22.4 of DSU speaks of equivalence not of reasonableness.

3.2.3 The Panel's Attitude Kills Two Instances of Adjudication

Korea might (or might not) have challenged the W-W methodology. We will never know. What we do know, though, is that it cannot do so. The case is over, as the Arbitrator, by using W-W as counterfactual, has immunized itself from challenges.

3.2.4 No Guarantee that Due Process Is Served

On the same wavelength, were this approach to be followed in other casesl, due process could be disserved. We explain. In this case, because the US authorities did calculate the dumping margin under all three methodologies (W-W; T-T; W-T), Korea had uninhibited access to the files. But, it could be the case that another investigating authority proceeds to calculation under W-T straight away. In this case, an Arbitrator, who assumes that a W-W dumping margin is the appropriate counterfactual, has, ipso facto, deprived the investigated entities from their due process

---

only one reasonable period of time during which implementation must occur, and clearly the measures discussed during the second recourse to a compliance panel have been taken after its end.

rights to question the consistency of the determination with the Antidumping Agreement, had the W-W methodology been privileged..

3.2.5 Are Panels Equipped to Choose the Reasonable Benchmark?

Panels are ill-equipped to perform the kind of review the Arbitrator has in mind. It works in *US–Washing Machines* because the calculated W-W margin was part of the record. But, what if it had not not? What is the guarantee that panels are knowledgeable enough to calculate reasonable dumping margins?[29] Even if we assume knowledge, what is the guarantee that panels have the information available to calculate reasonable dumping margins?

Or, would the Arbitrator have come up with a different counterfactual, in case W-W had not been available? But which counterfactual margin would emerge given that the Arbitrator itself asserted that Korea could not reasonably expect a 'no AD duties' scenario? By its own words, the Arbitrator had invited some calculation of the dumping margin, an exercise that we can seriously doubt they could complete without violating all of the points we have made so far.[30]

4. What To Do To End Tinkering?

The DSU regulation leaves a lot to be desired when it comes to discussing enforcement. This is probably one reason why the Arbitrator has behaved in erratic manner, as discussed above. In what follows, we advance a few thoughts to remedy this part of the overall enforcement process.

4.1 The Easy and Obvious Way Out

One quick and sure way to avoid problems is for the Arbitrator to always presume that the counterfactual is a state of the world where the illegality had not been committed at all. One might retort that our proposed solution might prove Draconian for minor violations. We have two responses to that.

First, there is a self-selection of disputes, and it is highly unlikely that complainants will challenge inconsequential violations. Indeed, the record so far of disputes by and large supports this view.

Second, in the unlikely case it happens again, minor violations can easily be addressed at the compliance stage, so that they provide no benchmark for calculating retaliation. The absence of retroactive remedies, *de facto* anyway, in the WTO system supports this view.

4.2 If We Were to Improve the Process through Legislative Amendment

Since we are at it, here is our proposals for amending the DSU, and improving the quality of the process.

4.2.1 Should the Original Panel Act as Arbitrator?

It does not make much sense to have the original panel act as Arbitrator. The original panel is too immersed in the dispute, and possibly subconsciously influenced by the process. The Arbitrator in *US–Washing Machines* was well immersed in the intricacies of the investigation before the

---

[29]Investigating the panel process, the conclusion in Mavroidis and Neven (2017) was that recourse to economics/econometric expertise is largely a function of the preferences of individual panelists, and not a matter of compulsion, whenever warranted.

[30]The USDOC dumping margin calculations can be based on thousands of individual transactions, or in some cases tens of thousands of transactions. While it might be conceivable that all confidential transaction data can be put on the record for a straightforward zeroing calculation, there is no assurance that in other disputes the parties would know what confidential data need to be submitted. Tens of thousands of pages of support information can be submitted as part of an antidumping of countervailing duty case. It is impossible for parties to know what part of their submissions will be deemed relevant for the Arbitrator's counterfactual.

USDOC, and did not hesitate to employ information from the files (the computer code output), without inquiring into the legitimacy of this approach.

A pair of fresh eyes, insulated from the original dispute, is needed to ensure that it will not use and, alas, abuse the factual record.

4.2.2 Appeals against Arbitrators' Reports

The evaluation of retaliation is probably the single most important aspect of the process. Enforcement, absent moral persuasion, largely hinges upon credibility of threat. Yes, the major stumbling block here is the statutory threshold in Article 22.4 of DSU, which amounts to cheap (temporary) exit from the contract. While we have sympathy for re-thinking the statutory threshold, a comprehensive discussion of this point escapes the ambit of this paper.

We do think, nevertheless, that the Arbitrator's decision should be appealable. Extending the process by a few months might be worth it, if this is the way to avoid tinkering with the counterfactual, and thus also reducing the potential for the errors we have discussed above.

4.2.3 Return to Suggestions

A more daring amendment would be to ask panels to issue binding suggestions, and thus provide interested parties with a ready-made counterfactual.

4.2.4 Rethink and Rewrite Article 22.6

Retaliation kicks in from the end of the reasonable period of time. But it is calculated based on damage inflicted years before. And here is the paradox. The WTO wants retaliation limited to the damage suffered, but not based on the damage suffered. For if that were the case, retaliation would have to retroactive, and not prospective.

By making retaliation prospective, we ask from Arbitrators to make generous assumptions. Think of it this way. Assume the total market size of the US washers' market between 2011 and 2014 (when Korea suffered damage) was $1 billion, and Korea's market share was 40%. Korea's loss was $400,000,000. The end of the reasonable period of time, let us assume, was 2017. And by that time the washers' market might be worth $0, because of a new invention, or $2 billion, because of a change in consumer preference. In either case, Korea will still be entitled to $400,000,000 no matter whether the market is booming or has gone bust.

What is then the purpose of retaliation? If it is to make sure that WTO members will be incentivized to perform their contractual obligations, the current regime achieves the exact opposite in case of booming markets. US has strong incentives to keep Korean washers out of the expanding US market, and would not mind paying the $400,000,000 to this effect.

5. Aftermath and Impact on Future Article 22.6 Disputes

Following the Arbitrator's award, the US authorities conducted a sunset review of the AD/CVD orders on Korea and Mexico. In that review the USITC determined

> that revocation of the countervailing duty order and the antidumping duty order on large residential washers ('LRWs') from Korea would not be likely to lead to continuation or recurrence of material injury to an industry in the United States within a reasonably foreseeable time.[31]

In its determination the USITC cited the decision by LG Electronics and Samsung to build large washer production facilities in the United States.[32] With the AD/CVD orders revoked, the United

---

[31]USITC (2019), p. 3.

[32]USITC (2019), p. 20 ('We find that LG and Samsung are likely to maintain their plans to supply the US market primarily from their new US washer production facilities after revocation.').

States is now deemed in compliance.[33] Critically, the US did not achieve compliance by imposing W-W margins without zeroing on Korea. In fact, as part of the sunset review proceedings the USDOC is required to report to the USITC what it expects the dumping margins to be if the case is sunset. USDOC reported that LG and Samsung would return to dumping at the same margins it originally calculated in 2011 – rates based on zeroing.

What are the other major lessons to be learned here? First, the critical importance of the Arbitrator decision that the N/I calculation did not have to be based on the termination of the order cannot be understated. This decision will likely have a profound impact on the N/I calculation in many future disputes. Conveniently, given that the issue was zeroing, the reasonable alternative duty could be easily computed in this case. Thus, the main challenge in this proceeding was the task of computing counterfactual trade values in the remedy year. The challenge of correctly (or reasonably) calculating the alternative duty in other disputes should not be minimized.

Second, the rejected US suggestion that information in the remedy year is all that is needed to compute N/I is dangerous to the WTO system. The US approach effectively incentivizes large violations as a large violation will drive trade to near zero, almost assuring a very small N/I.

Third, in a considerably more complicated dispute, *US–Anti-Dumping Methodologies (China)*, the Arbitrator implemented a very similar two-step approach. Many of the perils discussed in this paper are relevant to that dispute. To begin with, the WTO inconsistency at the core of that dispute ('single rate presumption') is profoundly difficult for the Arbitrator to model – both in scope (which firms are subject to the inconsistency) and in measuring the WTO consistent alternative duty. Nevertheless, the Arbitrator in that dispute essentially followed in the footsteps of what was done in *US–Washing Machines*.

References

Armington, P.S. (1969) 'A Theory of Demand for Products Distinguished by Place of Production', *IMF Staff Papers* 16(1), 159–178.

Bown, C.P. and R. Brewster (2017) '*US–COOL* Retaliation: The WTO's Article 22.6 Arbitration', *World Trade Review* 16, 371–394.

Bown, C.P. and M. Ruta (2010) 'The Economics of Permissible WTO Retaliation', in C.P. Bown and J.H.B. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*. Cambridge: Cambridge University Press, 149–197.

Francois, J.F. and K.H. Hall (1993) *COMPAS: Commercial Policy Analysis System*. Washington. DC, US: International Trade Commission.

Grossman, G.M. and P.C. Mavroidis (2007) 'Would've or Should've? Impaired Benefits Due to Copyright Infringement, US–Section 110(5)', in H. Horn and P.C. Mavroidis (eds.), *The WTO Case Law, The American Law Institute Reporters' Studies*. Cambridge: Cambridge University Press, 294–314.

Mavroidis, P.C. (2000) 'Remedies in the WTO Legal System: Between a Rock and a Hard Place', *European Journal of International Law* 11, 763–813.

Mavroidis, P.C. and D.J. Neven (2017) 'Land Rich and Cash Poor? The Reluctance of the WTO Dispute Settlement System to Entertain Economics Expertise: an Institutional Analysis', in M. Jansen, J. Pauwelyn, and T. Carpenter (eds.), *The Use of Economics in International Trade and Investment Disputes*. Cambridge: Cambridge University Press, 192–208.

Mavroidis, P.C. and T.J. Prusa (2018) 'Die Another Day: Zeroing in on Targeted Dumping – Did the AB Hit the Mark in US–Washing Machines?', *World Trade Review* 17, 239–264.

Messerlin, P. (2001) *Measuring the Costs of Protection in Europe: European Commercial Policy in the 2000s*. Washington, DC: Institute for International Economics.

USITC (2019) 'Certain Large Residential Washers from Korea and Mexico', Investigation Nos. 701-TA-488 and 731-TA-1199-1200 (Review), Publication 4882, April 2019.

---

[33]WTO Doc. WT/DSB/M/426 of May 20, 2019; WT/DS464/17/Add.17 of May 17, 2019.