2007

A Gambling Paradox: Why an Origin-Neutral 'Zero-Quota' Is Not a Quota Under GATS Article XVI

Donald H. Regan
University of Michigan Law School, donregan@umich.edu

Journal of World Trade 41(6): 1297–1317, 2007.

A *Gambling* Paradox: Why an Origin-Neutral "Zero-Quota" is Not a Quota Under GATS Article XVI

Donald H. REGAN*

In *US—Gambling*, the Appellate Body held that an origin-neutral prohibition on remote gambling (which is how they mostly viewed the United States law) was "in effect" a "zero-quota", and that such a "zero-quota" violated GATS Article XVI:2. That holding has been widely criticized, especially for what critics refer to as the Appellate Body's "effects test". This article argues that the Appellate Body's "in effect" analysis is not an "effects test" and is not the real problem. The real mistake is regarding a so-called "zero-quota" as a quota under Article XVI. That is inconsistent with the ordinary meaning of the word "quota" in such a context; it is at odds with the object and purpose of Article XVI; and it is likely to reduce Members' willingness to make concessions.

I. INTRODUCTION

In *United States—Gambling*,[1] the Appellate Body holds, or seems to hold, that an origin-neutral prohibition on the remote supply of gambling services is a "[numerical] quota" that violates GATS Article XVI:2(a) and XVI:2(c).[2] My object in this article is to explain why that is a misinterpretation of Article XVI. I do not discuss the other major issues in the case: whether the United States had actually made a commitment in respect of gambling; whether the Appellate Body was right to exclude from consideration all state laws; and the validity of the US Article XIV defense.[3] Also, to keep this relatively short, I assume the reader is familiar with the basic facts of the case and with the panel and Appellate Body opinions.

* William W. Bishop Jr Collegiate Professor of Law and Professor of Philosophy, University of Michigan. E-mail: <donregan@umich.edu>. For discussion and comments on drafts, I am grateful to Rob Howse, Markus Krajewski, Jürgen Kurtz, John Morijn, Federico Ortino and Joost Pauwelyn.

[1] *United States—Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/AB/R (adopted 20 April 2005).

[2] I refer to a "[numerical] quota" because Article XVI:2(a) speaks of "numerical quotas", while XVI:2(c) speaks of "quotas", unmodified. Whether this actually makes any difference is discussed in section II.C. below.

[3] I discuss the Article XIV defense in Donald Regan, *The Meaning of 'Necessary' in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing*, World Trade Review, November 2007.

II. GETTING THE ISSUE IN FOCUS

A. ORIGIN-NEUTRAL V. ORIGIN-SPECIFIC PROHIBITIONS

I say the Appellate Body "holds or seems to hold" that an origin-neutral prohibition on remote gambling is a market access violation. It is difficult to be certain just what the Appellate Body holds, because it gives no explicit attention (at least in this part of the opinion) to a distinction that is crucial for thinking about the case—the distinction between an origin-neutral (or facially non-discriminatory) prohibition on remote supply and an origin-specific (or facially discriminatory) prohibition. To make the distinction clear, and to open up the issue of its significance, consider a hypothetical country Home. For simplicity, we assume Home has a unitary government (no sub-central governments raising federalism issues); and we also assume it allows various forms of non-remote or face-to-face gambling, such as casino gambling. Now consider two possible legal situations with regard to remote gambling. (A) Home prohibits all remote supply of gambling services to consumers in Home, whether the supplier is domestic or foreign. This is an origin-neutral prohibition on remote supply. (B) Home prohibits remote supply of gambling services from abroad, but allows remote supply by domestic suppliers. This is an origin-specific prohibition on remote supply. I think it should be obvious that (A) and (B) raise different legal issues under GATS. Assuming Home has made an unlimited commitment on cross-border supply of gambling services, (B) is a paradigmatic violation of Article XVII. It is less clear whether (B) violates Article XVI, but it hardly matters what we say about that, given the clear Article XVII violation. In contrast, (A) is not a paradigmatic Article XVII violation; (A) violates XVII only if it is found that foreign remote suppliers are "like" domestic non-remote suppliers and that foreign suppliers are treated "less favorably". Nor is (A) a paradigmatic violation of Article XVI. I shall say more about why not later, but notice for now that even the *Gambling* panel concedes that (A) is not the sort of thing the drafters of Article XVI had before their minds.[4] My ultimate object in this article is to explain why (A) is not a violation of Article XVI at all. But for the moment, all I am trying to do is to get clear the logical distinction between (A) and (B), between an origin-neutral prohibition and an origin-specific prohibition.[5]

In its discussion of whether there is a market access violation, the Appellate Body makes no explicit reference to this distinction. One can easily suggest reasons why the

---

[4] *United States—Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/R (adopted as modified by the Appellate Body Report, 20 April 2005), para. 6.331, discussed in section IV.B. below.

[5] The actual *Gambling* case of course involves the United States, which is a federal country; and this is the source of some of the most vexing factual issues in the case. As I explain later, it is not part of my project in this article to say whether the overall United States scheme for regulating remote gambling did or did not amount to an origin-neutral prohibition; I am concerned only with whether origin-neutral prohibitions violate Article XVI. But lest the reader wonder, I regard it as clear that if the Wire Act, for example, forbids all cross-border and interstate remote gambling but allows intra-state remote gambling, that is not origin-neutral (unless it can be argued that state laws plug the gap by forbidding intra-state remote gambling).

Appellate Body ignores it. First, both (A) and (B) make all cross-border supply illegal. It is therefore very tempting, if one is trying to think in GATS terminology, to refer to both (A) and (B) as establishing a "zero-quota on cross-border supply", and then to frame the crucial question as whether such a "zero-quota on cross-border supply" is illegal. Both the panel and the Appellate Body seem to approach the matter this way. Unfortunately, proceeding this way suppresses the distinction between an origin-neutral prohibition on remote supply and an origin-specific prohibition, without any analysis of whether the distinction makes a difference under Article XVI. In further defense of the Appellate Body, it might be suggested that we already know Article XVI does not distinguish between origin-neutral and origin-specific measures; Article XVI clearly forbids origin-neutral quotas. True enough, but we must consider the possibility that an outright prohibition is not a "quota" in the relevant technical sense, even though it is undeniably tempting to refer to it as a "zero-quota" in this context. And we must consider the possibility that even though Article XVI forbids origin-neutral quotas, it does not forbid origin-neutral prohibitions.[6]

Although the Appellate Body does not explicitly address the distinction between origin-neutral and origin-specific prohibitions when it is discussing whether there is a market access violation, it is not unconscious of the distinction or wholly indifferent to it. The Appellate Body discusses whether the US law is origin-neutral or origin-specific in connection with the chapeau of Article XIV. It says, for example, "none of the three federal statutes [the Wire Act, the Travel Act, the Illegal Gambling Business Act] distinguishes, on its face, between domestic and foreign service suppliers".[7] And it is clear that this matters to it, because it upholds the Article XIV defense except in one particular: it rejects the defense insofar as the Interstate Horse-Racing Act may supersede those other Acts, and it does so precisely because it views the Interstate Horse-Racing Act as (at least possibly) facially discriminatory against foreign suppliers.[8] In other words, origin-specific prohibitions (or this particular origin-specific prohibition) violate the chapeau of Article XIV, while origin-neutral prohibitions do not. The fact that the Appellate Body generally upholds the Article XIV defense confirms that it views the US law on remote gambling as for the most part origin-neutral.

As to whether the US law really was origin-neutral, I express no opinion; the question is complicated and tangential to my project. My thesis in this article is not that the US prohibition was origin-neutral and therefore did not violate Article XVI. My

---

[6] Petros Mavroidis has argued that Article XVI does not forbid origin-neutral quotas at all, but only origin-specific quotas. *Highway XVI Re-Visited: The Road from Non-Discrimination to Market Access in GATS*, 6 World Trade Review 1 (2007), 1–23. Some of his arguments have considerable force, and I am not sure he is wrong. But the conventional wisdom is that XVI forbids origin-neutral quotas, and I accept that for purposes of argument precisely to highlight that even if XVI does forbid origin-neutral quotas, origin-neutral prohibitions are still different.

[7] *US—Gambling*, para. 354, cf. para. 357. In both of these paragraphs, the Appellate Body cites back to paragraphs in its market-access discussion that notably fail to provide support for the claim of origin-neutrality; but even so, that presumably indicates what the Appellate Body was thinking when it wrote the market-access section.

[8] Ibid., para. 369.

thesis is that *if* it was origin-neutral, it did not violate Article XVI.[9] The Appellate Body appears to have believed the US prohibition was (mostly) origin-neutral, and all of the formulations of its conclusions in the Article XVI discussion seem to apply to origin-neutral and origin-specific prohibitions indifferently. So the Appellate Body appears to have held that even an origin-neutral prohibition on remote supply violates Article XVI. If that is what it held, it was a mistake.[10]

B. WHY THE APPELLATE BODY'S "EFFECTS ANALYSIS" IS NOT THE PROBLEM

A number of other commentators have criticized the Appellate Body's finding of a market access violation.[11] They and I share the same general objection, that the Appellate Body's approach interferes too much with reasonable national regulation. But to my mind, their specific criticisms have been misguided. They have complained about an aspect of the Appellate Body's approach that is actually perfectly sound (which means their criticisms will not be persuasive to the Appellate Body), and they have missed the real mistake.

The Appellate Body's argument proceeds in two main stages. First, the Appellate Body argues that a total prohibition on cross-border supply is "in effect" a "zero-quota", at least in some loose, non-technical sense. Then it argues that this "zero-quota" is a quota in the technical sense of Article XVI:2(a) and XVI:2(c).[12] These two stages are not hermetically sealed off from each other, but the Appellate Body plainly recognizes that there are two distinct issues: (1) how far it can rely on the effects of a

---

[9] Some readers may regard the US law as clearly not origin-neutral. As I said in note 5 above, I agree that if the Wire Act, for example, forbids all cross-border and interstate remote gambling but allows intra-state remote gambling, that is not origin-neutral (unless it can be argued that state laws plug the gap by forbidding intra-state remote gambling). But I think the actual effect of the Wire Act—to say nothing of the relevant state laws—is less clear than most people assume.

[10] Not only was it a substantive mistake for both the panel and the Appellate Body to hold that an origin-neutral prohibition violated Article XVI; but this mistake is arguably the reason the litigation concerning implementation descended to something approaching farce. In the 21.5 proceeding before the panel, the United States argued that it had brought its invalid measures into compliance even though it had made no change to its law or administration. *United States—Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/RW (adopted 22 May 2007), paras 6.40, 6.57, 6.94. And the panel made a "pseudo-finding" on the facially discriminatory nature of the Wire Act, an issue it never decided the first time around, even though it acknowledged that it had no real jurisdiction to consider an issue the Appellate Body had apparently already decided the other way; seemingly the panel did this just in case the Appellate Body wanted to change its mind and needed some factual findings from the panel to point to; paras 6.118–6.123. All of this would have been avoided if in the original proceeding the panel had felt required to find facial discrimination as the ground for the Article XVI (or better Article XVII) violation. Sound law would lead to better-focused litigation.

[11] E.g, Markus Krajewski, *Playing by the Rules of the Game?*, 32 Legal Issues of Economic Integration 4 (2005), 417–447; Federico Ortino, *Treaty Interpretation and the WTO Appellate Body Report in US—Gambling: A Critique*, 9 Journal of International Economic Law 1 (2006), 117–148; Joost Pauwelyn, *Rien ne Va Plus? Distinguishing Domestic Regulation from Market Access in GATT and GATS*, 4 World Trade Review 2 (2005), 131–170; Joel Trachtman, *Decisions of the Appellate Body of the World Trade Organization*, 16 European Journal of International Law 4 (2005), 801.

[12] Article XVI:2(a) speaks of "numerical quotas"; in Article XVI:2(c), the word "quotas" occurs unmodified. It may seem that in this section I am sometimes careless about the possible distinction between "numerical quotas" and "quotas" simpliciter. Indeed, I am careless about the distinction here, and that is a considerable expository convenience. I explain in section II.C. below why I think the difference in language between XVI:2(a) and XVI:2(c) does not make any real difference, at least for our purposes.

measure for understanding the real purport of the measure, and (2) what sorts of measure, once properly understood, actually do violate Article XVI:2. Most of the criticism of the Appellate Body has focused on the first stage of the argument, on the consideration of effects in understanding the purport of the measure. The critics are worried that the Appellate Body might find, for example, that a licensing scheme for taxi drivers that requires them to show a high level of driving competence is "in effect" a forbidden numerical quota because it reduces the number of taxi drivers.[13] I agree this would be a serious misinterpretation of Article XVI, but if we attend to what the Appellate Body actually does in this first stage of the argument, there is no ground for worry that it would misapply Article XVI in this way. The problem, in my opinion, comes at the second stage, where the Appellate Body decides that an origin-neutral "zero-quota" is a quota in the technical sense. But other commentators have implicitly accepted this claim, or even explicitly endorsed it.[14]

Let me explain why I do not think the Appellate Body would find that the taxi-driver licensing scheme is "in effect" a quota. Even assuming that a stringent driving test reduces the number of taxi drivers below what it would otherwise be, there is a big difference between the effect of the remote gambling prohibition and the effect of the taxi-driver licensing scheme (both origin-neutral, I assume). Consider first the gambling prohibition. The prohibition does not name any number as the permitted number of remote suppliers, but it does unambiguously specify precisely how many remote suppliers there can be (zero), and we can infer this just from the face of the regulation. So it seems plausible to say the measure is "in effect" a "zero-quota" (so long as we understand this as a claim about what the measure amounts to, described non-technically, and not as a claim about its ultimate legality). If we object to this use of "effects" for understanding the nature of the gambling prohibition, we seem to be saying the gambling prohibition is not "in the form of [a] numerical quota" just because it does not use the word "zero". But to have such a hyper-formal criterion would invite evasion of the ban on numerical quotas. For example, it would entail that a law saying "there shall never be more remote gambling suppliers than are in business when this statute is adopted" is not a numerical quota, whereas it surely is (assuming there is at least one remote supplier when the law is adopted).[15] We simply cannot insist on formality to this extent.[16]

---

[13] This is Pauwelyn's example, in *Rien ne Va Plus?*, as note 11 above, at 160; it may seem a bit odd in connection with cross-border supply, but it makes the point. Ortino's example is a licensing exam for lawyers, in *Treaty Interpretation*, as note 11 above, at 141–142.

[14] Ortino suggests that the issue of whether a "zero-quota" is a quota has been settled (in the affirmative) by over 50 years of jurisprudence and practice under GATT. *Treaty Interpretation*, as note 11 above, at 134, n. 59. I explain in note 41 below, why I think the GATT jurisprudence does not settle the issue under GATS.

[15] Similarly, the Appellate Body points out, in its discussion of Article XVI:2(c), that if we say there can be no violation without an "express reference to numbered units", that would not cover "a limitation expressed as a percentage or described using words such as 'a majority'". *US—Gambling*, para. 250.

[16] To avoid any doubt on the reader's part, I would be happy to say that the gambling prohibition is equivalent "in effect" to a law that says in terms, "there shall be zero remote suppliers of gambling". But once we turn to the issue of legality, even that law should not be regarded as a quota in the sense of Article XVI (although I concede that any jurisdiction that wrote its law this way would be taking foolish chances with the interpretive process in dispute settlement). A "zero-quota" is not a quota at all, for reasons I explain in section III. below.

The taxi-driver scheme, in contrast, is not a quota, not even "in effect", because it does not specify the number of taxi drivers. We can see this just by asking, "What is the number? How many taxi drivers are allowed?" There is no answer. Even if the taxi-driver scheme reduces the number of drivers, it does not specify how many taxi drivers there shall be, nor does it put any upper bound on the possible number of drivers. Nor does it refer to a specific number of drivers even indirectly, as it might by saying that taxi licenses will be given to the top 10 percent of applicants at each administration of the test, or that there will be at most one taxi license per 1,000 population, or whatever. These arrangements might well be viewed as quotas "in effect", but they are still quite different from the driving test for taxi drivers.[17] In sum, the gambling prohibition "in effect" specifies precisely how many remote suppliers there shall be (zero), even though it does not name a number; in contrast, the taxi-driver scheme, although it presumably affects the number of drivers, does not specify a number, not even "in effect".

There is specific evidence in the Appellate Body's opinion that it would not regard the taxi-driver scheme as "in effect" a quota. It says the question of whether a measure is a numerical quota can be thought of as the question whether the measure is "quantitative" or "qualitative" in nature.[18] The Appellate Body could not possibly claim the taxi-driver scheme is "quantitative" as opposed to "qualitative" without completely eradicating the distinction. The scheme affects quantity, but only as a result of applying to individuals a qualitative test.[19] So the Appellate Body cannot regard the taxi-driver scheme as a quota.[20] Remember also, the Appellate Body says that although it is appropriate to look beyond the mere form of the measure (and we have seen it is right about that), "this is not to say that the words 'in the form of' should be ignored or replaced by the words 'that have the effect of'".[21] The Appellate Body sees that there must be some limit on the appeal to effects. Nothing in its opinion suggests to me that it would overstep that limit when characterizing measures in terms of what they do "in effect", if it had correct views about which measures, properly characterized, Article XVI forbids. The problem is not in the Appellate Body's consideration of "effects". The problem is elsewhere.

---

[17] I do not purport to settle here all questions about what should count as a quota. I am merely trying to explain why requiring taxi drivers to pass a special driving test obviously does not.

[18] *US—Gambling*, paras 225, 227, 232 (all stating that quotas are "quantitative") and especially paras 248, 250 (explicitly contrasting "quantitative" and "qualitative").

[19] One might suggest, under the influence of European Community law, that the licensing scheme is "in effect" a "quantitative restriction". And so it might be, under European law. But that is not a good model. The European Court of Justice has eradicated the distinction between "quantitative" and "qualitative" in this context. But the Appellate Body (working with different treaty language, we should remember) says it does not mean to.

[20] I explain later why I think the Appellate Body misapplies the "quantitative/qualitative" distinction when it decides that a "zero-quota" is quantitative. But I agree with it that the distinction is significant, and our present concern is with the taxi-driver case.

[21] *US—Gambling*, para. 232.

C. "NUMERICAL QUOTAS", "QUOTAS", AND "LIMITED-SCOPE QUOTAS"

GATS Article XVI:2(a), (b), and (d) all list as one sort of forbidden measure "numerical quotas"; Article XVI:2(c), in contrast, speaks simply of "quotas". Is there a real difference here? I think not, at least for our purposes. There are two ways to argue for this. One possibility is to argue that even XVI:2(c) is really about "numerical quotas". If one agrees with the United States (and the Appellate Body's assumption for purposes of argument)[22] that the phrase "expressed in terms of designated numerical units in the form of quotas" is a single descriptive item, as opposed to two separate items, then it refers only to quotas that are "expressed in terms of designated numerical units", and it is hard to see how such a quota could fail to be a "numerical quota", at least in effect. Indeed, one could argue from more general considerations that any measure that is properly regarded as a "quota" must be a "numerical quota", at least in effect. Quotas are about quantity, as the Appellate Body points out, and number is how we refer to quantity.[23]

But even though we can plausibly argue, both from the text and from general considerations, that XVI:2(c) is about "numerical quotas", I do not propose to rely on this argument. The reason is that if I first claim all of XVI:2(a), (b), (c), and (d) are about "numerical quotas" and then argue that a "zero-quota" is not a numerical quota, it might seem that my argument depends crucially on a claim that zero is not a number. In fact, I think one can make a good case that zero does not count as a "number" in this context,[24] but I do not want to proceed in a way that may seem to trivialize the argument and arouse the reader's suspicion that I am merely playing with words. I do not need to argue that zero is not a number. The reasons why a "zero-quota" is not a quota go much deeper than that.

So instead of assimilating XVI:2(c) to XVI:2(a), (b), and (d) by relying on the claim that every XVI:2(c) quota is a numerical quota, I assimilate XVI:2(a), (b), and (d) to XVI:2(c) by relying on the claim, which is unassailable in the present context, that every numerical quota must first be a quota. And instead of arguing merely that a "zero-quota" is not a numerical quota, I argue that it is not a quota, period. If this is true, then

---

[22] Ibid., para. 247.

[23] Someone might ask what we should make of a law that says, "The Gaming Commission may license remote gambling suppliers, but not too many." Is that a quota? Is it numerical? On the face of it, this law seems too unspecific to count as a "measure" at all. It has no content except what it is given by the Gaming Commission's practice. If we look at that practice, and if we find that the Gaming Commission is denying some licenses for reasons other than just the particular qualifications of the applicant, then it seems that the practice will almost certainly be viewable as creating either a "numerical quota" or a "monopoly" or an "exclusive supplier" arrangement or an "economic needs test", depending on just how the Commission explains its denials. Given the alternative possibilities just mentioned, I find it hard to imagine circumstances in which we would be inclined to say the practice created a quota (as opposed to a monopoly, or an economic needs test, or whatever) but not a numerical quota, not even in effect. So it still seems that any quota is a numerical quota.

[24] Zero is not among the most basic set of numbers that mathematicians work with, the positive integers. The positive integers are also known as the "counting numbers", and as I shall explain below, there is an intimate connection between quotas and counting. It took a long time in the history of mathematics for people to start thinking of zero as a number in the same sense as one, or two, or three; and their doing so was one step in a gradual distancing of the notion of "number" from our everyday intuitions, which continued, of course, with negative numbers, and irrational numbers, and imaginary numbers, and so on.

a "zero-quota" cannot violate the relevant portion of XVI:2(a), (b), (c), or (d). Hereafter I ignore the word "numerical", not because I have forgotten the wording of XVI:2(a), (b), and (d), but because I think the word "numerical" makes no difference to the present case.

One last preliminary matter about types of quotas. The panel expresses the view that a measure can be a quota even though it is what we might call a "limited-scope" quota: one that applies only to a certain method or methods of supplying a service, in which these methods do not cover an entire mode of supply.[25] For example, suppose the regulator imposes no restriction whatsoever on the supply of gambling services over the telephone, but imposes a limit of ten on the number of suppliers who may supply gambling services over the Internet. This limit on Internet suppliers is a "limited-scope" quota, covering only certain means of supply, but I agree with the panel that it should be regarded as a quota forbidden by Article XVI:2(a). My reason for calling attention to the proposition that a limited-scope quota is a quota is to point out that the panel never states this proposition separately, but runs it together with the distinct proposition that a "zero-quota" is a quota. For example, the panel says, in a finding that the Appellate Body quotes and upholds:

> [a prohibition on one, several or all means of delivery cross-border] is a "limitation on the number of service suppliers in the form of numerical quotas" within the meaning of Article XVI:2(a) because it totally prevents the use by service suppliers of one, or several or all means of delivery that are included in mode 1.[26]

As I say, this passage runs together the claim that a limited-scope quota may be a quota (which is correct) with the distinct claim that a "zero-quota" is a quota (which it is my object to argue against). We must not allow the second claim to borrow plausibility from association with the first.

III. Why an Origin-Neutral "Zero-Quota" Is Not a Quota Under Article XVI

Finally, we have isolated the crucial issue: is an origin-neutral prohibition, an origin-neutral "zero-quota", a quota under Article XVI? In arguing that an origin-neutral "zero-quota" is not a quota, I proceed in two stages. First, I explain why a "zero-quota" is clearly not a paradigm case of a quota. Even the *Gambling* panel acknowledges that a "zero-quota" is not the sort of measure the drafters of Article XVI had before their minds.[27] A "zero-quota" is a borderline case of a quota at best. But if it is a borderline case, how should we decide which side of the border it falls on? We should ask what is the point of listing quotas among the forbidden measures in Article

---

[25] E.g., *US—Gambling*, Panel, para. 6.338.

[26] *US—Gambling*, Appellate Body, para. 239, quoting Panel, para. 6.338. The Appellate Body notes that it was not asked to review the panel's reasoning on the status of limited-scope quotas.

[27] *US—Gambling*, Panel, para. 6.331, quoted approvingly in Appellate Body, para. 234. I explain in section IV.B. below why the panel's attempt to explain this on grounds other than a willingness to allow "zero-quotas" is implausible.

XVI—the point as revealed by consideration of the paradigm cases of quotas—and then we should ask whether that point would be served or disserved by forbidding "zero-quotas". The Appellate Body never considers the question this way, but I think it should have. So, the second stage of my argument is to explain why forbidding origin-neutral "zero-quotas" would not serve the point of Article XVI. (The case of origin-specific "zero-quotas" is more complicated.) After that, I turn in Part IV to the Appellate Body's arguments.

A. WHY A "ZERO-QUOTA" IS NOT A PARADIGM CASE

If we are asked to imagine a quota, what sorts of example come to mind? Perhaps an ordinance that says there shall be no more than 10,000 taxi medallions in some big city; or a regulation that says there shall be no more than 30,000 tons of tuna taken from a fishery in a particular year; or a law that says there shall be no more than one liquor license for every 2,000 inhabitants in any Michigan township. These are paradigm cases of quotas. In contrast, consider a law that makes it a crime to sell heroin. This is not, in ordinary speech, a "zero-quota" on heroin dealers or heroin sales. Rather, it is a prohibition. That is how we would normally refer to it. We would not even consider referring to a prohibition as a "zero-quota", except for Antigua's attempt to shoehorn a total prohibition into the categories of Article XVI:2.[28]

In sum, we normally think of a quota as a permitted number or amount greater than zero. This is the paradigm case. Something is allowed, or provided, or whatever, but only up to a certain number or amount. I shall refer to this sort of quota as an "n-quota" (where $n > 0$). A quota of one, or seven, or 43.5, or 17,000, is an n-quota. A "quota" of zero is not an n-quota; it is a zero-quota. (I have hitherto used quotation marks when referring to a "zero-quota", to emphasize that I did not think the word "quota" was really apt. This has probably got a little tiresome. Hereafter I refer to both n-quotas and zero-quotas without quotation marks.)

It may seem like a mere linguistic accident that we naturally use the word "quota" for an n-quota but not for a zero-quota. But it is not an accident. As is so often true, language is subtle, and it marks a fundamental difference (although a difference it is possible to overlook) in the way n-quotas and zero-quotas function. Suppose we have a law about remote gambling that says the Gaming Commission shall license not more than five remote gambling providers, each of whom must meet certain basic

[28] Revealingly, the claim that a prohibition is a quota seems to have been Antigua's third-line argument, after (1) the Article XVII argument, and (2) the claim that when an unlimited market access concession has been given in a particular mode, Article XVI:1 forbids any measure that interferes with access in that mode, without regard to the specific categories of XVI:2. Incidentally, there is one context I can think of in which we would naturally speak of a "zero-quota". Remember the fisheries regulation. Suppose that in successive years we have annual quotas of 30,000 tons, 17,000 tons, 23,000 tons, and then, after a weather disaster that temporarily causes a radical reduction in stocks, the regulator directs that no fish at all shall be taken in the next year. We could naturally call this a "zero-quota" for the year. But this usage of "zero-quota" is natural only because the quota is part of an ongoing program in which the quota varies from year to year and is usually not zero. There is no general prohibition on fishing, only a temporary prohibition.

qualifications with regard to freedom from criminal connections, financial soundness, website security, and so on. This is a standard n-quota (n = 5). Universal Gambling Amalgamated (UGA) now applies for a license. The Gaming Commission first determines that UGA meets the basic qualifications. But it still does not know whether a license should be issued. In order to know whether to issue a license, the Commission needs to know how many licenses have already been granted. If UGA has the basic qualifications, we cannot decide on UGA's case by looking at UGA alone; it is necessary to look at the cumulative results of other decisions (which might of course be the fact that no other previous decisions have been made, because UGA is the first applicant). This is the way an n-quota works. In contrast, if we have a zero-quota on remote gambling, which is to say, a flat prohibition, then we can always decide what to do about UGA's application without knowing how any other case has been decided. With a zero-quota, it is never necessary to look beyond UGA's application alone.

The reader might object that the reason we do not have to look beyond UGA's application to apply the zero-quota is that we know how many other licenses have been granted, at least if the law has been properly applied, namely, zero.[29] This misses the fundamental point that we do not have to look at the results in other cases because the results in other cases are irrelevant. But for the reader who thinks it is important that in the case of the zero-quota we know how many other licenses have been granted, I would point out that there is still a difference between the zero-quota and the n-quota. Under the n-quota, the only way to know how many other licenses have been granted is to count them. Under the zero-quota, we always know the number without counting.

I think that this indicates the proper way to understand the "quantitative"/"qualitative" distinction the Appellate Body advances to distinguish between quotas ("quantitative") and non-quotas ("qualitative"). In applying a true quota, such as an n-quota, we cannot decide the individual case before us (at least if the applicant is qualified) without counting the results in other cases. That is the sense in which a quota is "quantitative"; a question about the "quantity" of particular results in the whole ensemble of decided cases is essential to its application. By this criterion, the zero-quota is not quantitative; it is qualitative. We can apply the zero-quota to each case entirely on its own merits. We do not need to count; we don't need to think about quantity. Because the zero-quota is not quantitative, it is not a quota.

The Appellate Body relies implicitly on a different construal of the quantitative/qualitative distinction. They suggest that a zero-quota is "quantitative" because (unlike the taxi-driver testing scheme discussed earlier) it determines a specific number of licenses, namely zero.[30] Why is my construal of the quantitative/qualitative distinction, on which a zero-quota is not quantitative (and hence not a quota), better than their construal, on which it is? One reason is that my construal fits better with the ordinary

[29] In what follows, I continue to assume that any zero-quota is evenhandedly applied.
[30] *US—Gambling*, paras 227, 250.

usage of the word "quota", which we have seen does not include a zero-quota. But there is also an important non-linguistic reason. Even though a zero-quota does specify a particular number, the judgment behind the adoption of the zero-quota is still essentially qualitative. A zero-quota reflects a judgment that the service (perhaps further specified by mode or method) is somehow so objectionable or problematic that it should not be allowed at all. Such a judgment is essentially qualitative. Such a judgment does, to be sure, logically entail what quantity should be allowed, namely zero. But still, the fundamental judgment is qualitative (which explains, incidentally, why the zero-quota can be applied without counting). In the case of the n-quota, in contrast, where one has decided to allow some of the service, but also to impose some limit, then one must make an essentially quantitative judgment—how much to allow? Presumably a number of qualitative judgments about a variety of factors will feed into this quantitative judgment, but no qualitative judgment entails the quantitative conclusion in the same way that the judgment that a service is too problematic to be allowed at all entails the quantitative conclusion of a zero-quota. Setting the level of an n-quota requires a separate, essentially quantitative, decision (which also explains why, when one applies the n-quota, one has to count). In sum, the n-quota involves an irreducibly quantitative element; the zero-quota does not.

What have we accomplished so far? I think we have made a strong case that an n-quota is a quota, and a zero-quota is not.[31] But more important, we have demonstrated conclusively that a zero-quota is not a paradigm case of a quota; it is a borderline case at best. So it becomes necessary to ask whether the point of Article XVI is best served by holding that a zero-quota is a quota.

B. WHY AN ORIGIN-NEUTRAL ZERO-QUOTA IS NOT THE SORT OF MEASURE ARTICLE XVI AIMS AT

We need to focus the issue once again. Remember my overall thesis is that an origin-neutral prohibition, that is to say an origin-neutral zero-quota, is not a quota and does not violate Article XVI. In the previous section, I discussed zero-quotas without distinguishing between origin-neutral and origin-specific quotas. The arguments given there apply equally to both kinds of quota, and I think even an origin-specific zero-quota should not be thought of as a quota and should not be held to violate Article XVI.[32] Of course an origin-specific zero-quota is objectionable; but it is the origin-

[31] For the reader who remembers that according to the Appellate Body, the 1993 Scheduling Guidelines say a zero-quota is a quota, I explain in section IV.A. below why the Guidelines do not actually support the Appellate Body's position.

[32] There are two minor wrinkles in the way the arguments of the previous section apply to origin-specific zero-quotas. (1) I think it is clear that if there is no restriction on domestic suppliers, and a total prohibition (a "zero-quota") on foreign supply, it would never occur to us to call the prohibition on foreign supply a zero-quota in ordinary speech; but if there is an n-quota, say a 5-quota, on domestic suppliers, then we just might refer to the prohibition on foreign supply as a zero-quota. This is analogous to the case of the single-year zero-quota in a regulated fishery (see note 28 above), where the use of the word "quota" for the single-year prohibition is suggested only by the embedding of the prohibition in a larger scheme involving some true n-quotas. (2) The

specificity that makes it so. An origin-specific zero-quota is a core violation of Article XVII, and it should be dealt with under Article XVII.

But even though we did not need to distinguish between origin-neutral and origin-specific zero-quotas in the previous section, we do need to in this section. Once we ask whether forbidding zero-quotas would serve the same goals as forbidding n-quotas (which are the paradigm case under Article XVI), we find that origin-neutral zero-quotas and origin-specific zero-quotas are different. My main object in the rest of this section is to explain why it would not serve the goals of Article XVI to forbid origin-neutral zero-quotas. But I shall also explain briefly why it does serve the goals of Article XVI to forbid origin-specific zero-quotas (even if they are better dealt with under Article XVII). In practical effect, an origin-specific zero-quota is much more like an n-quota, even an origin-neutral n-quota, than it is like an origin-neutral zero-quota.

So, in order to decide whether an origin-neutral zero-quota is a quota, we need to know why n-quotas (including origin-neutral n-quotas) are forbidden by Article XVI:2.[33] (So that I do not have to keep repeating "origin-neutral", for the remainder of this article, when I talk about either zero-quotas or n-quotas, I shall mean origin-neutral zero-quotas or n-quotas, unless the context clearly indicates otherwise.) The most obvious suggestion about why n-quotas are forbidden is that they are likely to be protectionist.[34] If there is an n-quota on remote gambling service suppliers, for example, some firms are going to be allowed to supply remote gambling services and some are not. How is it decided which are the lucky ones? Almost any system, from grandfathering to a totally *de novo* selection process, is likely to favor local applicants. Even if the selection is *de novo* and evenhanded, local suppliers are likely to have the best current knowledge of the market, and thus to be the most likely to apply and to succeed. Any system that cuts off free entry over time seems likely to favor locals, as compared to a system which allows free entry and exit. The danger is that the restrictive system will be chosen because it favors locals.

A zero-quota does not pose the same danger of protectionism. If no one is allowed to supply gambling services remotely, then obviously there is no danger that the selection of who gets to do it will be skewed in favor of locals. There is not even *de facto* discrimination in favor of local remote suppliers. It might be argued that the total prohibition on remote supply is designed to protect local non-remote suppliers against competition from foreign remote suppliers. But that is just the sort of claim that should

---

cont.
origin-specific zero-quota is "qualitative" in the same sense as an origin-neutral zero-quota: it does not require us to count how many other licenses have been granted; and it specifies a number, zero (now for foreign supply), only as a logical consequence of a purely qualitative judgment—"we don't want foreign suppliers". But there is this difference: in the case of the origin-specific zero-quota, the qualitative judgment is itself objectionable to the general principles of the GATS and the WTO, unlike the origin-neutral qualitative judgment, "we don't want any remote gambling". This is why the origin-specific zero-quota is clearly illegal—but not because of its supposed quota-ness.

[33] Of course they are not "forbidden" at all if the regulating country has not scheduled the relevant sector, or if it has scheduled relevant limitations to its concession. But Article XVI clearly disfavors n-quotas. Why?

[34] Cf. Pauwelyn, *Rien ne Va Plus?*, as note 11 above, at 160.

be subjected to Article XVII's national treatment analysis, to consider whether remote and non-remote supply/suppliers are "like" services/suppliers and whether the regime treats foreign suppliers "less favorably". A zero-quota (origin-neutral, remember) cannot be used to protect domestic remote suppliers against foreign remote suppliers, but that is precisely what an n-quota threatens.

The objection to n-quotas might be broader than simply a fear of protectionism. Article XVI:2 might embody the view that n-quotas, even if they are not protectionist, do not usually make economic sense—they are unnecessarily trade-restrictive means for achieving whatever goals they might be aimed at.[35] I am not sure this is a more plausible explanation of Article XVI:2. But in any event, there is still a difference here between an n-quota and a zero-quota. The regulator that imposes an n-quota has decided that some amount of the relevant service (in whatever mode or method is specified by the measure, if it specifies) should be allowed. Seemingly, the service is not intrinsically problematic. We may of course still need licensing standards and health and safety regulations to guarantee that the service as actually provided is of the sort we regard as not problematic. But one might claim that once such standards are in place, it will generally be best to let the quantity supplied and the identity of the suppliers be determined by the market. That is, there is generally no justification for a quota. In contrast, the regulator that imposes a zero-quota has decided that the service (again, in whatever mode or method is specified by the measure, if it specifies) *is* intrinsically problematic. One cannot possibly argue that this sort of judgment is generally unjustified. So the zero-quota is still different from the n-quota.[36]

The title of Article XVI is "Market Access", so it is worth pointing out that my arguments about why a zero-quota differs from an n-quota depend on a fundamental difference in the way the n-quota and the zero-quota affect "market access". An n-quota says that with regard to some sector/subsector/mode/means of supply (as specified by the n-quota itself), some people will be allowed to supply that service and others will not. It is only because the n-quota allows supply by some people but excludes others that we worry about its being protectionist (in more than a *de facto* way that should be challenged under Article XVII); and it is only because the n-quota allows supply by some people but excludes qualified others that one might claim it is presumptively

[35] Cf. Ortino, *Treaty Interpretation*, as note 11 above, at 142–143.

[36] Rob Howse has suggested to me that a regulator that allows some of a service may still not regard that service as "not intrinsically problematic", as I claim in the text. For example, the regulator might want to stamp out brothels or betting-shops entirely, but knowing that if it tries it will only encourage illegal enterprises, it decides to license and supervise a limited number. This is a good point, and it suggests a reason why Article XVI:2 may be wrong to forbid n-quotas, at least without incorporating some "commercial purpose/non-commercial purpose" distinction. It does not suggest any reason why Article XVI:2, either as it is presently understood, or as revised or interpreted to deal with Howse's problem, should forbid zero-quotas. My object in this essay is not to say what an ideal Article XVI:2 would look like. It is generally assumed that Article XVI:2 flatly forbids n-quotas, and my object is to explain why even the best justification for that does not entail a limitation on zero-quotas. We may possibly need an overbroad ban on n-quotas to have a bright-line rule (subject always to exculpation under Article XIV). But there is no reason to include zero-quotas in this overbreadth; they are a bright-line category themselves (hence easy to exclude from the ban), and the possible justifications for a formalistic ban on n-quotas do not apply to zero-quotas at all.

unjustified. The zero-quota, in contrast, says no one will be allowed to supply the specified service (again, as specified by the zero-quota itself), and so it avoids these worries about the n-quota. One obvious way to describe this distinction between the n-quota and the zero-quota is to say that the former, but not the latter, denies some suppliers access to an existing market. With the n-quota, there is a market in operation for the specified service, but some people are denied access; some people are forbidden to do the very same thing the license-holders are allowed to do. With the zero-quota, no market in the specified service is permitted; there are no license-holders who are allowed to do what the suppliers who object to the zero-quota want to do. So there is no market that the objectors are denied access to.[37]

So far we have been comparing n-quotas and zero-quotas. Everything we have said about the distinction between them is confirmed if we consider the other sorts of measure that are forbidden by Article XVI:2(a)-(d)—monopolies, exclusive service suppliers, and economic needs tests.[38] All of these measures allow some suppliers to operate in the specified market, and prevent other qualified suppliers from doing the exact same thing the licensed suppliers are allowed to do.[39] They permit the market to exist, but exclude some would-be participants. Hence they can be abused for protectionist purposes, or they can plausibly be treated as presumptively unjustified. All of these measures are like n-quotas, and zero-quotas are still different from all of them.

That is to say (remembering my linguistic stipulation), *origin-neutral* zero-quotas are different from all of them. But notice that an origin-specific zero-quota is like an n-quota and the other forbidden measures. It allows the existence of a market, which is to be occupied only by domestic suppliers, and it forbids other, foreign, suppliers from doing the very same thing the domestic suppliers are allowed to do. As I said before, in practical effect, an origin-specific zero-quota is much more like an n-quota, even an origin-neutral n-quota, than it is like an origin-neutral zero-quota.

Some readers may still feel dissatisfied with my argument that Article XVI does not cover an origin-neutral zero-quota. If we state the argument crudely, it may seem that I am just saying an origin-neutral zero-quota is not a quota under Article XVI because it is not discriminatory. This may seem to blur the distinction between Article XVI and Article XVII, or to ignore the fact that Article XVI covers non-discriminatory quotas as

---

[37] I am suggesting that if the market is shut down completely, as by an origin-neutral zero-quota, there is no denial of market access, because there is no market. Some readers might object that shutting down the market completely is a denial of market access, the most extreme possible. I concede that in the abstract, either of these ways of thinking about when there is a denial of market access might make sense. But my suggestion makes it clear, at least, that there is a natural way to see the difference between n-quotas and zero-quotas in terms of whether they deny market access. And in fact the arguments we have advanced about why a zero-quota does not pose the same dangers as an n-quota are also arguments why in this context my interpretation of "market access denial" is the better interpretation. Excluding some suppliers from a market that is permitted to exist is problematic in ways that closing the market entirely is not.

[38] I ignore XVI:2 (e) and (f). What I say applies to them also, *mutatis mutandis*, but they are not so closely parallel as (a)–(d), and they do not speak of "quotas".

[39] Logically, something called an "economic needs test" might be used to close a market entirely, but in practice economic needs tests are always about protecting existing providers in some market against "destructive competition" from new entrants.

well as discriminatory quotas. Both objections are misguided. The difference between Article XVII and Article XVI is that Article XVII requires the issue of discrimination to be considered case by case, for each challenged measure (this is true whether the discrimination is *de jure*, and easy to identify, or *de facto*, and more problematic). In contrast, Article XVI, even if it is motivated by a worry about discrimination, does not require case-by-case consideration of whether the measure is discriminatory; it establishes a formal criterion. But with any formal criterion, it will sometimes happen that we are confronted with a borderline case that calls for refinement of our understanding of the criterion. Such a case is the origin-neutral zero-quota under the "[numerical] quota" criterion. To decide what to say about the origin-neutral zero-quota, we must ask what is the underlying justification for the criterion as it applies to the core cases, and the obvious suggestion is that the criterion is designed to forbid measures whose form justifies a presumption of discrimination.[40] My argument has been that there is no basis for such a presumption with regard to origin-neutral zero-quotas, so they should not be regarded as within the scope of the criterion. This does not blur the distinction between Article XVI and Article XVII, since it remains true that Article XVI involves a formal criterion (which we now understand better), and Article XVII requires case-by-case consideration of discrimination. Nor does this argument overlook the fact that Article XVI "covers non-discriminatory quotas". All we can say about Article XVI's coverage of non-discriminatory quotas without begging some presently relevant question is that Article XVI covers some quotas that are formally non-discriminatory, namely, origin-neutral n-quotas. But this is consistent both with the possibility that the underlying justification for the formal criterion is a worry about discrimination, and with the fact that origin-neutral zero-quotas do not trigger such a worry and so should not be covered.[41]

---

[40] We have also discussed the possibility that the formal criterion is designed to exclude measures whose form justifies a presumption of economic irrationality in some broader way. But that possibility is not relevant to the objections to my argument that I am now addressing.

[41] We now have the background to discuss Federico Ortino's suggestion that 50 years of GATT law, under which zero-quotas are illegal, has made it clear beyond question that a zero-quota is a quota; Ortino, *Treaty Interpretation*, as note 11 above, at 134, n. 59. This seems to me wrong on two counts. Most importantly, even if we concede for purposes of argument that a zero-quota is a quota under the GATT, that does not settle the question of whether a zero-quota is a quota under Article XVI of the GATS. We know the same word can mean different things in different contexts in the WTO agreements. (The classic example is the word "like", which the Appellate Body has told us means different things in two sections as close as GATT Article III:2 and III:4. *European Communities—Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001, para. 99.) And the context for the word "quota" is utterly different in the GATT and the GATS. Under GATT Articles XI-XIII, the quotas of interest are always origin-specific. They are quotas on foreign goods, or on goods from a particular country. But GATS Article XVI deals equally, indeed primarily, with origin-neutral quotas. (I say "primarily" because origin-specific quotas are forbidden independently by Article XVII.) The difference between origin-specific and origin-neutral is hugely significant when we are discussing zero-quotas, as I have explained in the text. So even if "quota" in the GATT includes zero-quotas, that is of minimal relevance to the question of whether an origin-neutral zero-quota is a "quota" under GATS Article XVI.

Now the other point. It may seem odd that I said we might "concede for purposes of argument" that a zero-quota is a quota under the GATT. Is not that much obvious? No, not at all. It is obvious that under GATT total bans—that is, "zero-quotas"—on imports and exports are illegal. But to say zero-quotas are illegal is not to say they are illegal as quotas. Both GATT Article XI:1 and XIII:1 speak specifically of "prohibitions", and zero-quotas are prohibitions (origin-neutral or origin-specific, and possibly of limited scope, as the case may be). So zero-quotas are illegal under GATT as the prohibitions they are; the question of whether they are illegal as quotas need never arise.

Let me summarize quickly the arguments why an origin-neutral zero-quota is not a quota under Article XVI.

(1) A zero-quota is not a "quota" according to the ordinary usage of that term in ordinary circumstances.
(2) Zero-quotas are essentially qualitative, and this, in conjunction with the Appellate Body's own premise that the crucial distinction is qualitative-versus-quantitative, entails that zero-quotas are not quotas.
(3) The only apparent reasons for disfavoring n-quotas (that they can be presumed to be protectionist or economically irrational) do not apply to origin-neutral zero-quotas.
(4) Origin-neutral zero-quotas do not deny any supplier access to an existing market for the precise service that supplier wants to provide.

In sum, an origin-neutral zero-quota, which is to say an origin-neutral prohibition of a mode or method of supply, is not an Article XVI quota.[42]

---

cont.
In contrast, GATS Article XVI does not talk about "prohibitions", which is why we find ourselves worrying about whether a prohibition can be called a "zero-quota" and subsumed under the word "quota". The crucial issue under GATS is simply not presented under GATT. (And of course, the reason GATT XI and XIII can talk about prohibitions, while GATS XVI does not, is that the former are limited to origin-specific measures.) I have not gone and looked at every GATT case that ever talked about quotas. I would be surprised to find much discussion of "zero-quotas". The one place where we might find such talk is where a country has a scheme that allocates an overall quota among other countries; and some country's zero-share is referred to as a "zero-quota". But that is like the case I mentioned in note 28 above, where we call the unusual single-year ban on fishing a "zero-quota" for that year. The use of the phrase "zero-quota" is parasitic on its being part of a general scheme where some years' quotas, or some countries' quotas, are not zero. In sum, there is no reason to think that in general a zero-quota (a prohibition) is a quota under GATT, even when it is clearly illegal.

[42] Let me mention, in order to reject it, a totally different argument from mine about why the United States remote gambling ban is not a market access violation. I think some people want to argue as follows: (a) the remote gambling ban is a "technical standard" in the sense of Article VI:4/5; (b) Article VI:4/5 and Article XVI are mutually exclusive; therefore, (c) the remote gambling ban does not fall under XVI. This argument seems to me misguided both in its general strategy and in substance. As to the strategy, I agree that we should try to make sense of all the provisions of GATS together, and I agree that that may affect the reading of individual provisions; but I think that when it comes to questions about the scope of XVI and VI:4/5, both provisions are clearer taken on their own (if we read them thoughtfully) than is any independently-sourced proposition about whether they are exclusive, or exhaustive, or whatever. So (b) in the argument above puts the cart before the horse. As to the substance, I think the claim that the remote gambling ban is a "technical standard" in the sense of Article VI:4/5 is both mistaken and pernicious. It is mistaken because "technical standard" must bear an analogy to "qualification requirement" and "licensing requirement", with which it is yoked in Article VI:4/5. The ban on remote gambling is not at all like a qualification or licensing requirement. This can be confirmed by looking at VI:4(a) and (b); the ban on remote gambling is not about "competence and ability to supply the service" or about "ensur[ing] the quality of the service" in any natural sense. (Cf. Panagiotis Delimatsis, *Don't Gamble with GATS—The Interaction between Articles VI, XVI, XVII, and XVIII GATS in the Light of the US—Gambling Case*, 40 Journal of World Trade 6 (December 2006), 1059–1080, at 1071, arguing that what we normally mean by a "technical standard" is a requirement that it is possible for suppliers to meet, at least in principle, thus gaining access to the market.) Some people may also be moved by the thought that Article VI is about "Domestic Regulation", and the remote gambling ban is a domestic regulation, so it must be covered by VI:4/5. This is a bad inference. Article VI:1/2, which address "all measures of general application affecting trade in services" arguably cover all domestic regulation (in committed sectors), but VI:4/5 explicitly distinguish themselves by specifying a narrower category of regulation that they cover (qualification requirements, technical standards, licensing requirements). As to why regarding the remote gambling ban as a "technical standard" covered by VI:4/5 is not only mistaken but pernicious: that is because it would authorize the Council for Trade in Services to adopt disciplines on regulations like the remote gambling ban under Article VI:4. There is nothing in the text to suggest this was the drafters' intention. In sum, the remote gambling ban falls neither under Article VI:4/5 nor under Article XVI:2 (as we have shown in the text). Some people seem to think all domestic regulation must fall under one or the other of these disciplines, but there is nothing in the text of the Agreement to suggest that.

IV. THE APPELLATE BODY'S ARGUMENTS: OF SCHEDULING GUIDELINES AND SCHEDULING STRATEGIES

A. THE IRRELEVANCE OF THE 1993 SCHEDULING GUIDELINES

The Appellate Body makes two arguments which we have not yet considered to show that an origin-neutral zero-quota is a "[numerical] quota" under Article XVI:2. The first argument relies on the 1993 Scheduling Guidelines:

> [The 1993 Scheduling Guidelines] set out an example of the type of limitation that falls within the scope of sub-paragraph (a) of Article XVI:2, ... That example is: "nationality requirements for suppliers of services (equivalent to zero quota)". This example confirms the view that measures equivalent to a zero quota fall within the scope of Article XVI:2(a).[43]

At this point in its opinion, the Appellate Body is not attending to the distinction between origin-neutral zero-quotas and origin-specific zero-quotas, but we have seen that that is a very significant distinction. And with that distinction in mind, we see that the passage from the Guidelines tells us nothing about how to regard an origin-neutral zero-quota. First, since the Guidelines refer to "nationality requirements for suppliers of services", they are talking about origin-specific zero-quotas. The Appellate Body over-generalizes when it (implicitly) reads the Guidelines to confirm that an origin-neutral zero-quota is a quota. Second, the Guidelines are not binding, and their evidential significance is weakened if we reflect that both the writers and readers of the Guidelines may have been careless about the precise basis of illegality for a sort of measure (an origin-specific zero-quota) that is so obviously illegal (because of Article XVII). Finally, although I disagree with the claim that even an origin-specific zero-quota is a quota, we have seen that if need be, I can accommodate that claim, because in its practical effect the origin-specific zero-quota is actually like the measures specifically listed in Article XVI:2, whereas the origin-neutral zero-quota is not. In sum, the Scheduling Guidelines are beside the point.

B. MISCONCEPTIONS ABOUT SCHEDULING STRATEGIES

The Appellate Body's final argument is embodied in this passage that they quote from the Panel Report:

> Paragraph (a) [of Article XVI:2] does not foresee a "zero quota" because paragraph (a) was not drafted to cover situations where a Member wants to maintain full limitations. If a Member wants to maintain a full prohibition, it is assumed that such a Member would not have scheduled such a sector or subsector and, therefore, would not need to schedule any limitation or measures pursuant to Article XVI: 2.[44]

---

[43] *US—Gambling*, para. 237, quoting 1993 Scheduling Guidelines, para. 6.
[44] *US—Gambling*, Appellate Body, para. 234, quoting Panel, para. 6.331.

The basic problem with this argument is that it begs the question.[45] The Appellate Body (adopting the panel's argument) is trying to explain away the fact that Article XVI:2(a) does not seem to address an (origin-neutral) zero-quota. It explains this by saying that a Member that wanted to have a zero-quota would not schedule the subsector in the first place. But why not? The implicit argument seems to be that a Member that wanted a zero-quota would not schedule the subsector because that would prevent it having the zero-quota it wants (at least absent a specific limitation). But that presupposes that a zero-quota is forbidden by XVI:2(a), which is the ultimate question. So, the Appellate Body's argument begs the question.

We can try to revise the argument to avoid the question begging. Instead of saying the Member that wants a zero-quota will not schedule the subsector because that would prevent it having the zero-quota (which begs the question), we could say the Member will not schedule the subsector, even if it is clear that that would not prevent the zero-quota, because there is no positive reason to do so (or at least no non-deceptive positive reason—I assume away deception for the moment). Now the implicit argument is that if other Members know the scheduling Member intends to have a zero-quota, they will value the scheduling of the subsector at naught. But that is just wrong. By scheduling gambling, Member M could be saying to the other Members:

> We all understand that even though I am scheduling gambling without limitation, I retain the right to forbid [remote] gambling entirely, because that would not violate Articles XVI or XVII. And for the time being I mean to forbid [remote] gambling entirely. But by scheduling gambling, I am agreeing that if I decide to abandon the total prohibition on [remote] gambling, I will not replace it with an n-quota, or discriminatory regulation, or any other measure that violates Articles XVI or XVII.[46]

This is a non-trivial commitment. It may not be all the other Members would like to have, but it could be very much worth having, especially if the other Members calculate that Member M will be forced by domestic political pressure to allow some [remote] gambling in the future. In that case, M's scheduling of gambling will impose real constraints on how that "some" must be allowed.[47]

---

[45] I pass over the fact that, as stated, the argument does not even apply to the instant case. The US measures under review do not constitute a "full prohibition" on a sector ["Recreational, Cultural, & Sporting Services"], or a subsector ["Other Recreational Services (except sporting)]", nor even on "Gambling" (nor even, in principle, on cross-border supply of gambling, which might occur by mail). The US measures do no more than cut off the principal means for one mode of supply of services in a "sub-subsector", namely cross-border supply of gambling services. Of course we could restate the panel's argument so that it claimed a Member would refuse to schedule a sector or subsector whenever it wanted to maintain a "full prohibition" on any particular means for a single mode of supply in a sub-subsector. But to my mind, this greatly reduces whatever limited plausibility the panel's argument may initially have. Waiving this point, I shall proceed in the text as if "Gambling" were a subsector and subject to a full prohibition, and I shall explain why the panel's argument fails even then.

[46] A word of explanation for the occurrences of "[remote]" in the imagined statement in the text: the argument works the same way with or without the limitation at certain points to remote gambling; I have inserted the bracketed "remote"s just for verisimilitude, since it was in fact only remote gambling that the United States proposed to eliminate entirely.

[47] The text is written as if M has a ban on [remote] gambling in place at the time it schedules gambling. That corresponds to the facts of *US—Gambling* as understood by the Appellate Body. But the basic argument works the same if M does not currently have a ban on [remote] gambling, but wants to retain the possibility of imposing such

As the next variant, suppose it is not clear one way or the other whether a zero-quota violates Article XVI. Now the Appellate Body's argument might be that a Member that wants to have a total ban on [remote] gambling, and that behaves carefully and cautiously, would not schedule gambling because of the uncertainty. As an empirical prediction about Members' behavior this is problematic—the US behavior that led to the *Gambling* case seems to be a counter-example. In such big, complicated negotiations, there will be failures of care and caution, even on the part of the most capable countries. But suppose we grant the empirical claim. What follows? It seems the Appellate Body now has an argument to show that a Member that wants a zero-quota would not schedule the subsector; and that could explain why the drafters of Article XVI:2(a) seemingly did not address zero-quotas. But there is a paradox in this suggestion: the suggested argument depends on uncertainty about whether zero-quotas violate Article XVI:2(a), which is recognized by the drafters; but if there is uncertainty, that is itself a reason for the drafters of Article XVI:2(a) to address the matter more clearly. But even waiving that point, and even if we now think we have an explanation of why Article XVI:2(a) might not address zero-quotas even if they really should be illegal, the explanation still does not give us any positive reason to think zero-quotas should be illegal. And I have argued in detail in earlier sections of this essay that there is no such reason.

This failure to suggest any positive reason for thinking zero-quotas should be illegal is a general defect of the "scheduling strategy" arguments we are now considering. But I have saved this observation until now because I think there is an undercurrent in the Panel's opinion, and perhaps even to some extent in the Appellate Body's opinion, that suggests that in the presence of uncertainty about the scope of XVI:2(a), we should resolve the uncertainty against the regulation, because that will favor trade. This approach of resolving uncertainty against the regulator is a bad idea in general (among other reasons, it has no basis in the Vienna Convention approach), and it is a particularly bad idea in the services context.

If a Member does not know exactly what might be forbidden at the boundaries of Article XVI or Article XVII, and/or does not know exactly what it might want to do in some subsector in the future, but if it is confident that it does not want to violate the clear commands of Articles XVI or XVII (now or in the future), then that Member might schedule the relevant subsector and take the chance that in the future, if it wants to do something in an area of uncertainty that it believes is legal, it will be able to persuade the dispute settlement tribunals of its view. But if the Member knows that any uncertainty about the scope of Articles XVI or XVII will be resolved against it, then it

---

cont.
a ban in the future. As long as it is clear that a total ban would not violate Article XVI as a zero-quota, M can afford to schedule gambling without limitation, and it may thereby be giving other Members something of value, even though it can impose a total ban if it chooses. Incidentally, Rob Howse has suggested another sort of example. M might offer a commitment on services which are currently excluded from the Agreement because they are supplied "in the exercise of governmental authority", in anticipation of a possible future privatization (complete or partial). Again, this might be of significant value to M's trading partners, even though it has no present effect.

will not schedule the subsector. (Uncertainty about its own future desires will prevent it from scheduling the subsector with specific limitations.) And not having scheduled the subsector, the Member might actually end up later restricting trade in a way it would have been willing *ex ante* to agree not to do. Even if it never ends up doing anything it could not have done if it had scheduled the subsector, still the fact that it did not schedule the subsector will have reduced the predictability of its behavior, with consequent economic costs. In sum, it is a mistake to have a rule that gratuitously discourages concessions. This is especially so with regard to services, where there are still so many valuable concessions to be made. Rulings that seem to promote trade in the short run, by reading concessions expansively, may damage trade in the longer run, by discouraging concessions.[48]

But, finally, is it not unfair for the United States to schedule gambling, with no limitation in respect of cross-border supply, and then enforce a law that makes cross-border supply impossible? The short answer is, no, it is not unfair, if the law in question does not violate Article XVI or XVII. There is nothing unfair in Members' pursuing their own agendas within the limits of the treaty. Even an "unlimited commitment" in a sector is not a commitment to allow totally untrammeled supply of services; it is a commitment not to adopt or enforce measures that violate the specific terms of Articles XVI or XVII (or VI). Any other view would make the listing of specific violations in XVI:2 otiose.

Antigua may find itself disappointed if it thought all along that a zero-quota would be illegal. But if it is Antigua that has made the wrong guess about the appropriate legal result, it is Antigua that the decision should disappoint. Notice that in the face of uncertainty, Antigua could have attempted to secure its expectation by asking for a specific commitment not to have an origin-neutral zero-quota under Article XVIII. Similarly, if worried by uncertainty, the United States could have secured its expectation that it could have a zero-quota by entering a specific limitation on its commitment. In sum, if either side was not prepared to take its chances on the interpretation of unclear provisions, each side had something they could do about it.

There is one last possibility. If in the circumstances of a particular case we really think one Member has been unfairly surprised by the effect on its trade of another Member's measure that is legal under the Agreement, there is always the non-violation remedy under GATS Article XXIII:3. But even if the complaining Member prevails, the respondent Member will not be required to withdraw the measure. Presumably this remedy will be used with circumspection, as it is under GATT XXIII:1(b). And in the actual *Gambling* case, it seems very unlikely that Antigua was unfairly surprised relative to any reasonable expectations it had in 1994.

[48] Cf. Krajewski, *Playing by the Rules*, as note 11 above, at 437.

V. SUMMARY

An origin-neutral zero-quota differs both from an origin-neutral n-quota, which is the paradigm violation of the "[numerical] quota" language of Article XVI:2(a)-(d), and from an origin-specific zero-quota, if that is thought to violate Article XVI as well as Article XVII. The origin-neutral zero-quota is not protectionist; it is not presumptively economically irrational; it does not exclude anyone from an existing market in which other suppliers are allowed to do what the excluded supplier wants to do; it is based ultimately on a purely qualitative judgment. The argument that a Member that wants to enforce an origin-neutral zero-quota over a whole sector/subsector would not schedule that sector/subsector is not defensible; in appropriate circumstances it could make perfect sense to schedule, without limitation, a sector/subsector in which one wanted to enforce a zero-quota. Finally, there is no general unfairness in scheduling a sector/subsector and then enforcing an origin-neutral zero-quota. For the rare case in which it might actually seem unfair, the proper recourse is the non-violation remedy under Article XXIII:3.