

Audience Cost Theory and Its Audiences

BRANISLAV L. SLANTCHEV

I really wanted to read Trachtenberg's article as taking part in the "modeling dialogue" by bringing historical evidence to bear as a corrective to some of the assumptions in audience cost theory (ACT). There is much value in this kind of dialogue between theory and data, and any viable specification of ACT must take into account several of the issues he discusses.¹ Unfortunately, Trachtenberg opts instead for a broad indictment of scholars who work on the "purely abstract level" and who easily "get carried away with a particular idea" so that they "lose sight of all the reasons the particular effect . . . might not count for as much in the real world."² Neither the logic of Trachtenberg's arguments nor the historical evidence he presents warrants such a harsh conclusion.

Trachtenberg argues that:

(T-1) democratic leaders do not generate audience costs on purpose, or, more usefully:

(T-1a) it is democratic leaders who can use audience costs to tie their hands, and

(T-1b) they prefer not to do so and thus avoid making any statements that could potentially generate such costs;

(T-2) even when these leaders do commit, their opponents ignore them; and

(T-3) other factors better explain crisis outcomes.

In this necessarily terse note, I will show the following. Since ACT is silent about the link between regime type and ability to generate audience costs, (T-1a) actually criticizes something the theory does not say. In fact,

Branislav L. Slantchev is professor in the Department of Political Science, University of California, San Diego.

¹ Marc Trachtenberg, "Audience Costs: An Historical Analysis," *Security Studies* 21, no.1 (2012). On the modeling dialogue, see Roger B. Myerson, "On the Value of Game Theory in Social Science," *Rationality and Society* 4, no.1 (1992). On the modeling enterprise, see Robert Powell, *In the Shadow of Power: States and Strategies in International Politics* (Princeton, NJ: Princeton University Press, 1999), 24–29.

² Trachtenberg, "Audience Costs," 42.

in several of the cases Trachtenberg studies, one would not even expect ACT's mechanism to be at work for the simple reason that both sides had eschewed the use of force. Since (T-1b) follows only if audience costs really did tie the hands of leaders—otherwise why expend effort avoiding them?—Trachtenberg implicitly relies on one of the core assumptions of ACT. This is problematic because it concedes to ACT more than what theorists would and because it creates serious logical problems when combined with (T-2). If public commitments are hands-tying for leader A, then why would leader B disregard them? Conversely, if B disregards such commitments because they are worthless, then why would A try to avoid making them? Setting aside the problem of reconciling such claims, the evidence Trachtenberg presents to back up (T-2) is unpersuasive. Finally, I believe that Trachtenberg is correct about (T-3) for many crises. I will not discuss this further except to note that this point has been made before by scholars working on the theory.³ Overall, then, Trachtenberg's article fails as a critique of ACT although it is a welcome contribution to the modeling dialogue.

WHAT DOES AUDIENCE COST THEORY SAY?

It is perhaps useful to begin with a brief summary of what Fearon's original specification of ACT actually says.⁴ The brief discussion should make it clear that what often passes for ACT in empirical studies that investigate how regime type is linked to credibility is in fact no theory at all. For our purposes, the original ACT can be summarized as follows:

If (1) backing down in a crisis makes an actor suffer costs in addition to those arising from conceding the stakes, (2) these costs increase as the crisis escalates, (3) these costs can become so large that war becomes preferable to a concession, (4) no other mechanism for coercing the opponent exists, and (5) attempting to coerce the opponent does not increase his costs of conceding, *then* escalation can commit an actor to fighting, and the resulting risk of war discourages bluffing, which makes escalation informative and gives it a coercive role.

Setting aside requirements specific to the game-theoretic model and some obvious domain limitations—for example, a world with two actors and no future interaction—we are left with a list of at least five premises, all of which are assumed to be true and all of which might be necessary for the conclusions to hold. Let us now turn to Trachtenberg's arguments.

³ Kenneth A. Schultz, *Democracy and Coercive Diplomacy* (Cambridge: Cambridge University Press, 2001); Branislav L. Slantchev, *Military Threats: The Costs of Coercion and the Price of Peace* (Cambridge: Cambridge University Press, 2011).

⁴ James D. Fearon, "Domestic Political Audiences and the Escalation of International Disputes," *American Political Science Review* 88 no. 3 (September 1994), 577–92.

Observe that there is nothing in the summary above that connects the type of political system to crisis dynamics. The original theory is silent on whether democracies are more or less able to commit credibly during a crisis. This is what Fearon calls a “plausible working hypothesis.”⁵ Even though Trachtenberg acknowledges this in passing, he nevertheless insists that the idea is “taken quite seriously in the international relations literature.”⁶ This may be so, but one can hardly fault the theory for its misapplication. In fact, scholars, many working on the theory itself, have repeatedly emphasized that the notion of “democratic credibility through audience costs,” while indeed plausible, turns out to be deeply problematic.⁷ Thus, the thrust of Trachtenberg’s criticism (T-1a) is entirely misplaced.

Turning now back to what the theory does say, the first three assumptions constitute the core of ACT. They are critical but without micro-foundations.⁸ Why would an actor suffer costs for escalating and backing down? Fearon justifies the assumption with an appeal to *domestic* political opponents who would punish such behavior because they “deplore that *international* loss of credibility, face, or honor.”⁹ Following this cue, most attempts to rationalize audience costs have tended to conceptualize audience costs as punishments of leaders who “fail to fulfill their commitments” (for example, by failing to carry out a threat). The problem for ACT is to specify a mechanism such that audiences find it optimal to punish leaders who back down and leaders pursue strategies that result in such punishments in equilibrium. This is not the place to rehash some of the assumptions needed to get all ducks in a row. Suffice it to say, to get leaders, political opponents, and audiences to generate audience costs is quite difficult when they behave strategically and pursue their interests. Thus, scholars working out the theory have puzzled over its very foundations almost from the outset. Many have noted serious difficulties with the assumption that backing down in a crisis should generate any audience costs.¹⁰

⁵ Ibid., 582.

⁶ Trachtenberg, “Audience Costs,” 4.

⁷ Bruce Bueno de Mesquita et al., *The Logic of Political Survival* (Cambridge, MA: M.I.T. Press, 2003); Branislav L. Slantchev, “Politicians, the Media, and Domestic Audience Costs,” *International Studies Quarterly* 50, no. 2 (2006); Jessica Weeks, “Autocratic Audience Costs: Regime Type and Signaling Resolve,” *International Organization* 62, no. 1 (2008); Jessica Weiss, “Autocratic Signaling, Mass Audiences, and Nationalist Protest in China,” *International Organization* (forthcoming).

⁸ Alastair Smith, “International Crises and Domestic Politics,” *American Political Science Review* 92, no. 3 (1998); Kenneth A. Schultz, “Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War,” *International Organization* 53, no. 2 (1999); Slantchev, “Politicians, the Media, and Domestic Audience Costs.”

⁹ Fearon, “Domestic Political Audiences,” 581.

¹⁰ I will not dwell on the third assumption except to point out that Fearon is explicit about it. Ibid., 584. I have both noted how important it is for the results and doubted its empirical plausibility: “both [domestic audience costs and reputation] provide leaders with levers that they barely have access to in practice. As a consequence, the ability to commit derived from such devices is suspect.” Slantchev, *Military Threats*, 51, 61.

This work is important because in his attempt to refute the theory with (T-1b), Trachtenberg actually has to assume its fundamental mechanism—the “black-box” at the core of ACT that scholars have tried to open—is correct. Trachtenberg essentially claims that (democratic) leaders do not strategically attempt to generate audience costs because they do not want to be constrained by domestic opinion in foreign policy and because they fear the adverse effect explicit threats might have on the opponent or on third parties. Policymakers evade audience costs by making vague threats in which plausible deniability enables them to obfuscate whether they have failed to follow through on any particular commitments by finessing outcomes to the point that it is unclear whether they have made any concessions at all and by trying not to go public unless they are forced to. All of this means that audience costs should play little or no role at all provided leaders can get away with secrecy or are sufficiently able to mold public opinion.

I have no serious quarrel with this line of reasoning, but one should realize just how much of ACT’s core mechanism one has to assume to make it work. Leaders can only be constrained by public opinion if making clear commitments and then backing down from them does generate audience costs. The premise of this argument is thus equivalent to the first core assumption of ACT. Moreover, the conclusion that leaders would not generate such costs does not necessarily follow because leaders might not be able to deal in secret or manipulate public opinion. Trachtenberg provides evidence for both possibilities. For example, he notes that during the Fashoda crisis Salisbury was not a “free agent” and that he was forced to go public in order to show a belligerent public that he was not giving away the farm. Even though he did not deliberately seek to “stoke the public’s outrage,” the end effect was just the same, and “even minor concessions could not now be made.”¹¹ Trachtenberg also observes that even the autocratic Russian government could be “carried away by or unable to resist strong nationalist and pan-Slav feeling” during the Eastern Crisis of 1878, which caused leaders to refuse concessions until they found themselves isolated diplomatically and on the verge of war with Britain.¹²

Trachtenberg also provides direct evidence of leaders deliberately stoking public outrage, as Kiderlen did in the Second Moroccan Crisis of 1911. Kiderlen’s threats did arouse nationalist feeling in Germany, and the government “ended up paying a huge political price for what was seen as its willingness to accept a humiliating defeat,” as ACT assumes it would.¹³ Thus, we have evidence that leaders might not always avoid being constrained by their domestic audiences even if they wanted to, and we have evidence that

¹¹ Trachtenberg, “Audience Costs,” 15.

¹² *Ibid.*, 9.

¹³ *Ibid.*, 20.

sometimes they deliberately seek to inflame public opinion. Moreover, we have evidence that such constraints might occur in non-democracies as well, which is consistent with ACT as it stands now even if it does go contrary to the original “working hypothesis.”

The desire to avoid the loss of flexibility in foreign policy perhaps is not be the only reason policymakers might be loath to make overt threats. Frightening the opponent might backfire if it raises his audience costs, provokes him into attacking, or gets third parties involved in the dispute to the detriment of the threatener. Trachtenberg provides evidence for all of these concerns, and it does constitute omitted and potentially distorting factors in the original ACT. But it is not as if theorists are ignorant of this: Kurizaki explicitly deals with the possibility of a public threat increasing the cost of concessions for the opponent, thereby analyzing what happens when one relaxes the fifth assumption, and I show that if a clear threat can provoke countermeasures that worsen the threatener’s military position, the threatener might avoid making such threats.¹⁴

Thus, I agree with Trachtenberg that “public threat-making . . . is not necessarily an effective instrument of statecraft,” even though he seems to be unaware of the work that deals with these issues and at any rate does not note just how dependent on ACT’s fundamental structure many of these arguments, including critique (T-1b), are.

THE REACTION OF OPPONENTS

Whatever the points of agreement might be, Trachtenberg actually deemphasizes the first argument in favor of the second: even when (democratic) leaders go public—whether by design or not—adversaries do not infer that their hands are tied.¹⁵ Given the author’s own stress on this, it is very disappointing that it is precisely here where the evidence presented is the weakest. Consider first the consistency of critiques (T-1b) and (T-2) when taken together. If (T-1b) is correct and leaders try to avoid incurring audience costs because doing so would act as a commitment device, then it is unclear why their opponents would fail to understand when such a commitment does take place, as (T-2) would have it. To sustain such an argument, one would have to assume that these opponents are consistently and persistently dense. If, on the other hand, one grants (T-2) and assumes that opponents correctly ignore public posturing because it does not generate any commitments, then one must be puzzled by (T-1b), which insists that leaders actively try to avoid such posturing precisely because it is committing. Trachtenberg does not

¹⁴ Shuhei Kurizaki, “Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy,” *American Political Science Review* 101, no. 3 (2007); Branislav L. Slantchev, “Feigning Weakness,” *International Organization* 64, no. 3 (2010).

¹⁵ Trachtenberg, “Audience Costs,” 7.

attempt to reconcile these contradictory premises, but the weakness of that approach is evident when one tries to parse the arguments he makes in some of the historical cases and the evidence he presents for his interpretations.

Consider the Eastern Crisis, where Trachtenberg is quick to dismiss the importance of the British government going public. We are told that the Disraeli government did not go public in order to limit its freedom of action but to “make it clear . . . that Britain would go to war if Russia did not moderate her position.”¹⁶ This type of statement should really give one pause: even if we were to grant the unsubstantiated assertion that British leaders were committed to war in the absence of concessions, we still have two critical questions. First, how many concessions would they need in order to be induced not to fight? Second, how were they going to signal this—that is, “make it clear”—to the Russians? Since any concessions the Russians would agree to would be related to what they believed the British would be satisfied with (and thus depended on their estimate of the latter’s commitment to war), this must become a matter of signaling. It just will not do to assert, with scant evidence, that the Russians “did not doubt” the British commitment. What really matters is the extent of that commitment because, as Trachtenberg himself notes, Disraeli would have only resorted to force “if the terms were not satisfactory.” Thus, the British government did face the problem of convincing the Russians that they would have to give up more than they had originally believed. As to whether Disraeli’s threats were vague or not, I am unpersuaded by a sole reference to a speech Disraeli gave over a year before the crisis even began.

The only evidence for why British public statements did not matter for the Russians comes from the Russian ambassador Shuvalov who apparently advised his government not to be provoked by Britain’s bellicose statements and warlike public opinion.¹⁷ This is not the same as denying that this opinion might make it difficult for the British government to compromise. At any rate, Trachtenberg shows no evidence about what the Russian government thought of this. It might be instructive, however, that after the capitulation at Berlin, Shuvalov’s career was ruined because he was blamed for the Russian humiliation. Also troubling is that while Trachtenberg cites as evidence Shuvalov’s apparent lack of concern about British audience costs, he dismisses Gromyko’s concern about American audience costs in 1962.¹⁸

¹⁶ Ibid., 11-12.

¹⁷ Ibid., 11.

¹⁸ Ibid., 31. The lack of evidentiary support is also present in other cases (for example, the Second Moroccan Crisis). In other cases (for example, the First Moroccan Crisis), the mechanism is not supposed to apply since neither side was pursuing tactics that involved a risk of war, an essential ingredient to a coercive strategy of this type. James D. Fearon, “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs,” *Journal of Conflict Resolution* 41, no. 1 (1997); Slantchev, *Military Threats*.

On balance, then, it appears to me that there is more support for (T-1b) and not for Trachtenberg's preferred critique (T-2), suggesting that the problem of mutual consistency might not be critical.

CONCLUSION

Fearon's original model that introduced the notion of "audience costs" is both less than what supporters of audience cost theory think and more than what critics imagine. It is less because it does not specify the mechanism that would generate escalating audience costs but only studies their coercive effect assuming that they exist. It is also less because it makes some assumptions that can be shown to be distorting: it ignores private communication and other instruments of coercion. It is also less because its most popular conclusion—that democracies are "more credible" in crises—actually comes from unmodeled assumptions that connect regime type to the assumption about audience costs and is thus doubly removed from having firm theoretical foundations.

The model, however, is also more than what critics allege because it is not wedded to a particular substantive interpretation of its parameters, and as such, the fact that it assumes escalating audience costs makes the conclusion about their coercive effect independent of the generating mechanism: *if* such a mechanism can be found, then the conclusions automatically follow. It is more because it spurred research to provide such a mechanism by connecting domestic politics to crisis behavior. It is also more because it provided the foundation for the study of reasons leaders might try to avoid public commitments, opting instead for vague threats, and attempt to finesse crisis outcomes. Even if all these activities undermine the effectiveness of the audience cost mechanism as a coercive device, the fact that leaders try to avoid incurring such costs suggests a real concern that they might (which is consistent with the theory) and furthermore suggests that the theory's conclusions might be valid in those cases when leaders are unable to avoid these costs. This has sparked research on the conditions that can prevent leaders from freely manipulating public opinion. All of this just goes to show that the explicative and generative usefulness of Fearon's original insight has been staggering, not simply in the number of references, but in the development of new ideas in the context provided by the audience cost theory.