

Politics as if Evidence Mattered: A Reply to Achen and Bartels

Anthony Fowler and Andrew B. Hall

August 2018

Christopher H. Achen and Larry M. Bartels's response to our study indicates more consensus than one might think. They agree with us that shark attacks do not, in general, lead voters to punish incumbents, stating that they "would not expect" the "typical shark attack" to affect a presidential election. This consensus is important since many readers have thought that their claims were stronger and more general than they are.

The only remaining point of disagreement is over the purported electoral effects of shark attacks in New Jersey in 1916. The evidence offered consists of two regressions with 20 and 14 non-independent observations, respectively, for which the results depend upon specification choices, a historical error, using the aberrant 1912 election to control for baseline preferences, and incorrect standard errors.

We agree with Achen and Bartels that historical context and substantive knowledge are vital for studying politics, but none of the historical facts and conjectures presented in their reply constitute persuasive evidence that shark attacks influenced elections even in this one case. Their claim that we ignore historical evidence is especially unpersuasive since it was our historical research that uncovered consequential errors in their analyses.

As we discuss below, we think this exchange is an important chance for the field to think about what constitutes persuasive evidence for an empirical claim. While quibbling over shark attacks may seem like academic navel-gazing, Achen and Bartels's claim underpins a large scholarly and public discussion over the health of democracy. Our field must determine the standards of evidence it expects for claims of this magnitude, and we hope that this memo, which briefly responds to several important points Achen and Bartels raise in their reply, contributes to this ongoing discussion.

Consensus: Shark Attacks Do Not Generally Influence Elections

In their reply, Achen and Bartels agree that shark attacks do not affect elections, aside from one instance in New Jersey in 1916. They go on to say that their argument about irrelevant events affecting elections should only apply when there are "substantial economic costs in the affected communities, significant publicity, and explicit attributions of blame to the president or the federal government."

If there are significant costs and explicit attributions of blame, then perhaps we should question whether we're really sure the event is politically irrelevant. If Achen and Bartels's position is that voters only blame politicians for irrelevant events in the rare cases that they blame politicians, then they're begging the question, and they haven't presented a falsifiable, scientific hypothesis. And even with these stringent conditions that Achen and Bartels lay out, we suspect they could have just as easily described locust swarms and Spanish flu as meeting these conditions had those events appeared to influence elections—but those results were null.

In any event, Achen and Bartels's discussion on this matter represents important progress for this literature, which has generally interpreted the shark-attack finding as an example of a general phenomenon. The subtitle to their influential working paper was "Why Shark Attacks are Bad for Democracy." In the paper, they described their goal as studying "the impact of shark attacks on American presidential elections." Other scholars have clearly interpreted the claim to be general. Writing in *Pacific Standard*, Seth Masket states that "voters punish their leaders for . . . shark attacks." In his review of Achen and Bartels's book in the *Journal of Politics*, Neil Malhotra writes that "voters frequently punish incumbents for things they cannot control such as shark attacks." Achen and Bartels have now clarified that their claim is specific to only the shark attacks in New Jersey in 1916. Future

discussions of “blind retrospection” should take note of this new consensus. Rather than claiming that shark attacks indicate a general failure of electoral accountability, Achen and Bartels say that voters *do not* blame incumbents for shark attacks in about 99 percent of all recorded shark attacks in American history.

Did Shark Attacks Affect Voting in New Jersey in 1916?

Agreeing that shark attacks outside New Jersey in 1916 do not affect elections, Achen and Bartels go on to defend their evidence in that particular case. Our critique of their evidence takes several different forms.

First, we show that their original results are sensitive to various coding and specification choices. Essentially, our analyses suggest that Achen and Bartels may have inadvertently obtained a false positive result amidst a large set of specification choices that all could have been justified *ex post* had that specification been the one to produce the desired result. They respond to these concerns in several ways. First, they provide rationalizations for their preferred specifications. And second, they show, in some cases, that their specifications fit the data better (as measured by statistics like the r-squared). Unfortunately, neither of these approaches compellingly distinguishes a false positive from a genuine result. Of course, many different specifications can be justified and rationalized *ex post* (see [here](#)). And selecting a preferred specification by maximizing fit is essentially inadvertent p-hacking.

Second, we show that the 1912 election, which Achen and Bartels use to control for the baseline political preferences of counties and towns, was anomalous. Figure 3 of our paper clearly shows that their county-level result is driven by the unusualness of 1912, not anything that happened in 1916. This is not particularly surprising since 1912 was one of the most interesting and unusual presidential elections in U.S. history—with three presidents on the ballot and with the incumbent Republican coming in third place. Curiously, despite their emphasis on the importance of historical context, Achen and Bartels never address or acknowledge this issue in their reply.

Third, we show that Achen and Bartels’ standard errors are misleading. If we apply their inferential strategy to state-elections with no shark attacks, we detect an effect as large as theirs 32 percent of the time, and the estimate is statistically significant ($p < .05$) 27 percent of the time. Achen and Bartels respond by saying that “No student of American politics should be surprised to discover that Texas counties along the Gulf of Mexico are different in many ways from those inland or that the politics of New York City and Long Island are distinct from those of upstate New York,” but that’s precisely our point. Coastal regions are different from non-coastal regions and they are subject to different idiosyncratic shocks in the absence of shark attacks. Therefore, when we see that coastal counties trended differently in one instance in New Jersey, we don’t know whether that was the result of shark attacks or whether that’s the kind of thing that happens from time to time regardless of shark attacks.

What’s the Matter with Seaside Park?

Much of Achen and Bartels’s reply concerns the single town of Seaside Park. They originally excluded Seaside Park from their town-level analysis in Ocean County, stating that the town “apparently split into two between 1912 and 1916.” In our study, we combed through historical records of town boundaries and discovered that Seaside Park did not, in fact, split in two or change boundaries over this period. Achen and Bartels’s original justification for excluding Seaside Park was erroneous.

Confronted with this historical error, Achen and Bartels now provide a new and completely different explanation for why dropping Seaside Park was the right thing to do. Specifically, Achen and

Bartels speculate that the residents of what is now Seaside Heights illegally cast ballots in Seaside Park in 1912 but not in 1916 after they had their own township and polling location.¹

We have two responses to this claim. First, the evidence is entirely circumstantial. Achen and Bartels claim that it would have been difficult for the residents of what is now Seaside Heights to vote in their assigned polling location on the mainland and therefore they must have voted illegally in another town. That logical leap is not supported by any direct evidence. Between September of 2016 and January of 2017, we communicated with Christopher J. Vaz, a historian of Seaside Heights, who told us he is aware of no evidence that the residents of what is now Seaside Heights voted in Seaside Park in or before 1912. Furthermore, he tells us that a regular boat service connected Seaside Heights to the mainland and that “it was common for Seaside Heights residents to use the boat service to travel to the mainland for shopping, doctor and dentist visits, and business related meetings downtown.”

Second, even if some residents of Seaside Heights did vote illegally in Seaside Park in 1912, it’s clear that Achen and Bartels did not apply the same standard when deciding which other towns to include or exclude. They only appear to have looked closely into this issue after we pointed out that their initial justification was erroneous, perhaps because we showed that the exclusion of Seaside Park is essential for producing their preferred result.

Town-Level Analysis of All Four Beach Counties Shows Precise Non-Effect

As part of our town-level analyses, we note that it’s surprising that Achen and Bartels only focus on Ocean County—just one of their four beach counties. We attempt to replicate their analyses pooling all four beach counties, and we obtain a precise null result. Puzzlingly, Achen and Bartels only discuss this briefly in a footnote, and they attempt to dismiss this evidence only by raising a series of questions. Here are our answers to those questions.

“To ensure meaningful comparisons, did they drop areas whose populations grew substantially?” We explain this in our paper. The answer is *no*, although, we can use the same 25 percent rule as Achen and Bartels and the result does not change meaningfully (the coefficient goes from .004 to -.003 and the p-value goes from .703 to .839).

“Similarly, were they able to closely match beach and nonbeach townships and boroughs on prior voting behavior?” We’re not exactly sure what they’re asking here since Achen and Bartels never made an explicit effort to match the beach and nonbeach towns in Ocean County. Pooling all beach counties, the beach townships differed from the nonbeach towns by about 5 percentage points in 1912, which is about the same difference that we see in Ocean County alone.

“Did they have the same difficulties distinguishing classic Jersey Shore communities from non-Shore areas that they report having in Ocean County?” Yes, towns are not neatly divided into beach and nonbeach territory in either Ocean County or the other beach counties. Following Achen and Bartels’s coding decisions in Ocean County, we attempted to be conservative in coding beach towns, only including those towns that are virtually entirely beach communities.

“What did they do about the Garden State Parkway?” We use that as the dividing line for all counties and we include all towns in our analysis that have some territory to the east of it, with the exception of Lacey, which Achen and Bartels exclude.

“What did they do about Neptune Township?” Neptune has both beach and near-beach territory, so consistent with the coding rules stated above and in the paper (and consistent with Achen and Bartels’s implied recommendations), we code this as a nonbeach town.

¹ While Achen and Bartels are, of course, correct that election administration was quite different in the early 20th century, voting outside one’s assigned polling place was a crime in New Jersey in 1912 (see “[An Act to Regulate Elections \(Revision of 1898\)](#)”). Though enforcement may have been uneven, the penalties were serious, including a \$1,000 fee (equivalent to roughly \$25,000 today) and up to 5 years in prison.

Historical Context is Important, and Our Study Was Informed by It

A major theme of Achen and Bartels's reply is that historical context is important and that we ignore it. At various points, they claim that we "[set] aside the politics and history of the Progressive period," that we engage in "presentism," that we "proceed by ignoring or overriding...historical evidence," and that we lack "careful engagement with the relevant theory and evidence." We wholeheartedly agree with them about the value of historical context, substantive details, and carefully combining qualitative and quantitative evidence. But we reject their view that our enthusiasm for rigorous quantitative evidence has somehow led us to neglect history or context. Indeed, it was precisely our focus on history and context that led us to uncover Achen and Bartels's historical error regarding Seaside Park and to discover that their results depend on the historically aberrant 1912 election.

Why Should We Care About Shark Attacks?

Presumably, few living people care *per se* about whether shark attacks influenced voting in New Jersey in 1916. So why should anyone care about this debate? First, rightly or wrongly, bold and broad claims about voter competence and the health of democracy—by Achen and Bartels and others—have been based on this narrow claim that shark attacks influenced voting behavior in this one unusual case. Given those claims, we should know whether the empirical finding is reliable or not.

Second, beyond this particular question, our field is at an important crossroads. Presumably, everyone who has bothered to read this exchange cares deeply about what counts as persuasive evidence of a political phenomenon. Quantitative social scientists have been complacent or poorly informed for too long about forking paths, incorrect standard errors, and false positives. And that complacency appears to be particularly acute when the finding in question is newsworthy or aligns with our predispositions. We need to decide as a field how seriously we take evidence and how much we care about uncovering the truth as opposed to producing exciting findings.

So how could one generate compelling evidence of an ex-ante unlikely phenomenon like shark attacks influencing elections? Obviously, strong claims require strong evidence, and we all have to decide where we draw the line regarding our willingness to believe false claims or miss out on exciting new findings. Preregistration is one strategy for reducing false positives, but it's not a viable option for studies using data that have been publicly available for over a century. With a small number of observations and with correlated shocks between beach and non-beach areas, the effective standard errors associated with these kinds of regressions are huge—meaning few findings of this particular sort can be persuasive. But if multiple independent sources of evidence all point in the same direction, claims like those about shark attacks and elections can become persuasive. For example, if the point estimate from New Jersey in 1916 were bigger than virtually all of the placebo estimates, it would be harder to explain this result away as a chance false positive. If the Ocean County result had held up in the other beach counties, that would be additional persuasive evidence of a genuine phenomenon. If the aggregate analyses using many elections and shark attacks had shown clear evidence of an effect, we would have been even more persuaded. And if those estimated effects were stronger in situations where we would theoretically expect the biggest effects (e.g., election years, presidents seeking reelection, multiple fatalities), the case would be genuinely convincing. However, we find no compelling evidence that shark attacks influenced voting in New Jersey in 1916 or in any other case.