

# Expectation Setting and Retrospective Voting

**Neil Malhotra** Stanford University  
**Yotam Margalit** Tel Aviv University

*That citizens engage in retrospective voting is widely established in the literature. But to what extent is retrospection affected by the expectations that leaders set in advance? We develop a theoretical framework of how expectation setting affects voters' retrospective evaluations of incumbent performance. To test the theory, we conduct a series of between-subjects experiments in which we independently manipulate both expectation setting and the eventual outcome. In domains where politicians have practical authority, or direct influence over outcomes, setting high expectations incurs a cost in public support if the projected outcome is not attained. The same is true in domains where politicians have theoretical authority, or limited influence, but where expectation setting sends a signal about the leader's judgment. However, in domains where politicians have neither practical nor theoretical authority, setting high expectations is unambiguously beneficial, implying that optimism is valued by voters as a personality disposition.*

One of the most replicated and stable empirical regularities in political science is that voters punish politicians for poor performance and reward them for good performance (e.g., Fiorina 1981; Key 1966; Kramer 1971; for a review, see Healy and Malhotra 2013).<sup>1</sup> Such electoral behavior is often touted as promoting democratic accountability by sanctioning failure and promoting the selection of more competent leaders (Fearon 1999; Ferejohn 1986). Recent research has made significant headway in exploring the microfoundations of retrospective voting behavior (e.g., Fernández-Albertos, Kuo, and Balcells 2013; Huber, Hill, and Lenz 2012; Margalit 2011; Tucker 2006). Yet despite the extensive literature on retrospective voting, a surprisingly understudied question is how expectation setting conditions performance evaluation. While it is clear that the reelection prospects of incumbents are harmed by bad outcomes, to what extent is retrospection affected by the expectations set in advance by the leader? Is the electoral impact more severe if beforehand the leader set high expectations by expressing optimism about the eventual outcome? Put more formally, should models predicting incumbent reelection include not

only indicators of performance but also measures of prior *expectations* of performance?

Recent examples underscore the importance of understanding how expectation setting affects retrospective voting. In order to gain passage of the American Recovery and Reinvestment Act (ARRA) of 2009 (commonly known as “The Stimulus”), President Barack Obama’s economic team presented a graph suggesting that the ARRA’s passage would help lower the unemployment rate to 5.4% by the end of 2012. Republican candidate Mitt Romney (2012) pounced on these optimistic projections in the presidential debates, noting that despite the passage of the stimulus plan, unemployment remained persistently high: “The president said by now we’d be at 5.4 percent unemployment. We’re 9 million jobs short of that.” The counterfactual question is: Would President Obama’s reelection prospects have been brighter if he had managed expectations differently by projecting that the unemployment rate would fall to about 8%, the eventual level at the end of his first term, rather than the more optimistic estimate. Similarly, in the debate regarding a possible invasion of Iraq in 2003, Vice President Dick Cheney (2003) advocated American action by predicting that the U.S.

<sup>1</sup>An online appendix for this article is available at <http://dx.doi.org/10.1017/S0022381614000577> containing supplemental analyses. Data and supporting materials necessary to reproduce the numerical results will be made available at Neil Malhotra’s IQSS Dataverse upon publication.

forces “will, in fact, be greeted as liberators” by the Iraqi civilians. As it turned out, U.S. forces met persistent and violent resistance by locals for many months after the invasion, a point repeatedly emphasized by President Bush’s critics. Indeed, the consequences of the war, particularly the local incidence of fatalities, led to a significant drop in President Bush’s vote share in the 2004 presidential election (Karol and Miguel 2007). Would the political impact of the Iraq War have been any different if prior to military action the President and his team managed expectations differently by predicting a more somber (and exacting) unfolding of events?

Whether politicians get punished electorally for having been *too* optimistic (i.e., for setting lofty expectations on which they did not deliver) is an open question. Yet the view that voters prefer candidates who they perceive to be “sunny optimists” is almost a political convention. For instance, in their coverage of the 2012 presidential election, *The Los Angeles Times* noted that “In presidential elections, voters often head toward the more optimistic candidate, and Obama at least for now has put himself in position to grab that advantage” (Hennessey and Parsons 2012). Similarly, *The New York Times* claimed that “one axiom of politics is that the optimistic candidate wins . . .” (Stolberg 2011), and *The Daily Telegraph* noted that Obama “was top of the list in terms of being seen as the most optimistic candidate—a key sign of electability” (Harnden 2007). This conventional wisdom has also received some scholarly support. Zullo and Seligman (1990) conducted an extensive textual analysis of over 80 years of presidential candidates’ election speeches and found that the candidate who was more of a “pessimistic ruminator” lost in the large majority of the cases.

What then is the role of expectation setting in shaping voters’ assessments of government performance? One possibility is that voters respond positively to high expectations only when the eventual outcome is positive, in which case optimism could be mistakenly seen as a valued disposition, whereas what voters may actually reward is the politician achieving her stated goals. Alternatively, voters may assign value to an optimistic disposition in and of itself and reward elected officials for setting a hopeful target irrespective of the eventual outcome. Casual analysis typically adopts this second perspective but does not disentangle it from the first. Unfortunately, we have little evidence to distinguish between these different conjectures. While there is scattered evidence that voter expectations of policy performance

are consequential (e.g., Kimball and Patterson 1997; Waterman, Jenkins-Smith, and Silva 1999), there exists no research on whether the targets set and projections made by politicians affect how voters judge their subsequent performances in office.

We argue that the key to understanding the moderating relationship between expectation setting and retrospective assessments is to distinguish policy domains in their degree to which politicians are perceived to influence outcomes. Voters value optimism as a character trait or a personality disposition they find appealing. Yet voters do not value optimism as a substitute for performance. Consequently, in those domains where elected officials influence the outcome, we contend that optimistic projections will harm politicians if the outcome ultimately turns out to be bad. Similarly, even in domains where politicians do not have meaningful influence but their expectation setting signals judgment, voters will punish optimistic officials if events do not fall in line with projections. In contrast, in domains perceived to be beyond the politicians’ control and where no judgment signal can be extracted, voters reward high expectation setting even if the projection is wrong and the outcome turns out to be negative. We construct a simple formalization of this logic, which guides the set of between-subjects survey experiments we designed to test the argument.

Across six studies conducted in two democratic countries investigating multiple domains and outcome variables, we prompt several thousand participants with short accounts of elected officials making various types of projections—some reflecting high expectations, some not—and provide information about the eventual outcome. The accounts deal with a set of domains (the economy, security, the environment, weather/tourism, sports) that vary in terms of: (1) the degree to which the politician has plausible influence over the outcome and (2) whether a meaningful signal about the leader’s judgment can be extracted by voters. By randomly assigning projections and outcomes, we can causally estimate the moderating effect of expectation setting on retrospective performance evaluation. Further, we replicate our experiments multiple times to ensure that the results are stable and consistent.

Most directly, this article contributes to the study of retrospective voting. The robust finding that voters respond to government performance is a cornerstone of scholarly work on electoral behavior. Yet a striking feature in much of this research is that the incumbent is conceived as an actor whose “input” into the retrospective voting decision solely consists of government

performance. This study highlights the role of expectation setting as an additional—and substantively consequential—way incumbents can influence voters' assessments. More broadly, the results also add to our understanding of the functioning of the democratic process. If voters value and expect (even unconsciously) that politicians will be unabashedly optimistic, it can disincentivize campaigns from dealing with hard issues that might require appearing pessimistic.

Relatedly, the economic voting literature has increasingly explored how a country's exposure to the globalized economy affects the relationship between performance of the national economy and voters' attitudes toward the ruling government (e.g., Hellwig 2008). The key claim of this line of research is that voters punish incumbents to a lesser extent for negative conditions when the government's control over the local economy is perceived as weaker. Our argument is different. We stress the role of theoretical authority, which suggests that even when leaders are perceived to have limited control over a certain domain, in this case the local economy, voters might nonetheless punish the incumbent for a disappointing outcome if lofty expectations turn out to be incorrect.

We first present some stylized facts suggesting that optimism and high expectation setting are traits that the public recognizes and values in political leaders. Next, we present a theoretical model of expectation setting and voter responsiveness to motivate and assist with our experimental designs. The subsequent section describes our empirical approach and experimental setup. We then present the results and assess their broader implications for the study of retrospective voting.

## Do Voters Value High Expectation Setting?

Before describing the experiments that are the core of our empirical analysis, we present some stylized facts that suggest that people conceive of optimism and high expectation setting as desired and potentially valuable traits among politicians.<sup>2</sup> We asked a national sample of Americans: "Generally speaking, how optimistic or pessimistic do you want political leaders in the United States to be? Try not to think of any specific office holder. Instead, please

<sup>2</sup>All data in this section come from Study 2, described in further detail below.

tell us what you would like to see in politicians generally," presenting them with a 7-point scale ranging from "extremely pessimistic" (1) to "extremely optimistic" (7). The mean response to the question was 5.3, located well above the neutral midpoint category of "neither pessimistic nor optimistic" and between "slightly" and "moderately" optimistic. Notably, this preference for optimistic leaders does not simply seem to reflect people's own optimistic nature. We also asked respondents: "Some people expect future events to be positive; such people are called 'optimists.' Other people expect the future to be negative; such people are called 'pessimists.' How would you rate yourself on the following scale?" When individuals rated themselves on the same scale of this character trait, the mean response was 4.5, exactly between the neutral option and being "slightly" optimistic. Examining responses across different groups of respondents reveals remarkable consistency of the pattern.<sup>3</sup> While members of some groups are more likely to describe themselves as optimistic (e.g., the religious, Democrats), the overall pattern is clear: individuals tend to locate themselves in the slightly positive range and want their leaders to be even *more* optimistic than they are.

To get some sense of how people perceive optimism and pessimism as compared to other potential traits of leaders, we also asked respondents to rank a set of six traits that a politician may possess: "optimistic," "pessimistic," "overly optimistic," "overly pessimistic," and two additional traits that were randomly selected from a set of six: "cautious," "courageous," "intelligent," "daring," "charismatic," and "reliable."<sup>4</sup> The order in which the six traits were presented was fully randomized to ensure that primacy effects did not bias the rankings. As before, we find that people assign a high value to optimism as a leadership quality, and that premium is consistent across different groups in the electorate (see online Appendix 2). In fact, optimism is on average ranked as a highly preferred trait among leaders (average rank: 2.2), above qualities such as being cautious (2.6), charismatic (3.0), or daring (3.0), and about as important as being courageous (2.1). The only two

<sup>3</sup>Detailed results can be found in online Appendix 1.

<sup>4</sup>The question wording was: "In supporting or opposing political candidates, people tend to care about different personality traits. Please rank the following traits of a possible candidate from *most positive* (1) to *least positive* (5). In other words, place what you think is the best trait at the top."

traits ranked as more important are reliability (1.4) and intelligence (1.5).<sup>5</sup>

These patterns indicate that, at a minimum, individuals *think* that they prefer optimistic politicians in the abstract. We next explored whether optimism was valued when examining actual politicians. We conducted a LexisNexis search to collect all media reports from U.S. presidential elections held over the past three decades which mentioned a presidential candidate's name.<sup>6</sup> We then coded the number of articles in the 12 months leading up to the election in which a presidential candidate was discussed in the context of optimism. Second, we asked respondents in the survey to report their retrospective impressions of the same politicians. Respondents were asked to rate the optimism of each candidate on the same 7-point scale described earlier, where a "7" denotes an "extremely optimistic" candidate, and the midpoint refers to a candidate who is "neither optimistic nor pessimistic." Figure 1 highlights the correlation between the portrayal of candidates in the media and voters' perceptions of the candidates' optimism. Most notable, and perhaps least surprising, is the fact that the two most popular presidents of the past 40 years—Ronald Reagan and Bill Clinton—were also those most frequently described in media reports in the year before the election as optimistic and those remembered by survey respondents as such. The view of Clinton as optimistic by both the media and survey respondents is particularly striking when contrasted with his successor as the Democratic Party's candidate for president, Vice President Al Gore. As the figure indicates, Gore was rarely described as an optimistic individual, a fact that was made all the more acute given the more frequent mention of his opponent George W. Bush as optimistic.<sup>7</sup>

These stylized facts suggest that there may be a connection between candidates' perceived nature as optimists and their appeal among voters. Yet these illustrated patterns do not demonstrate causality; one cannot conclude that candidates were electorally successful because of their (perceived) optimism.

<sup>5</sup>Also notable is the fact that optimism is not only ranked highly, but that even being "overly optimistic" (4.2) is seen as a more desirable trait than being pessimistic (4.5); in contrast, being "overly pessimistic" is, by far, the least appreciated quality in the list (5.6). Nonetheless, people do penalize politicians for being overly optimistic, a point we return to below.

<sup>6</sup>Details of the procedures for this content analysis can be found in online Appendix 3.

<sup>7</sup>The media mentions of Obama's optimism is likely biased downward due to the normalization procedure used (see online Appendix 3).

Whether and under what conditions voters reward high expectation setting among elected officials is an open question which we turn to address in the subsequent sections.

## Theoretical Overview

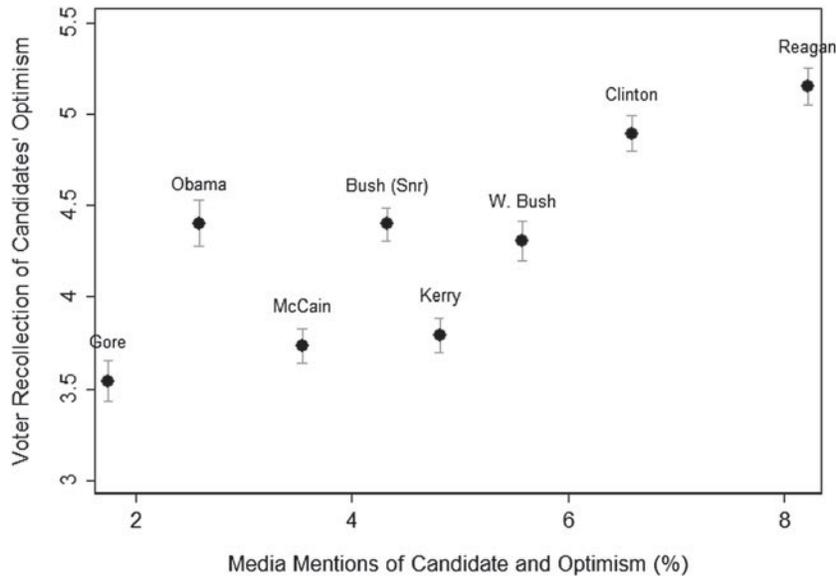
In this section, we present a simple decision theoretic framework for a politician's decision to set a high versus a low expectation. The goal here is not to lay out a thorough formal theory but rather to produce a tractable framework regarding a narrow set of parameters to guide the experimental design.

When voters observe an outcome on which an incumbent politician had set an expectation a priori, the comparison of the predicted outcome to the actual outcome can affect voters' assessment of the incumbent in several ways. Our model focuses on three factors. First, voters learn whether the incumbent's prediction was correct. If the leader assessed the available information and reached an accurate prediction, that might boost voters' confidence in the leader's judgment. Second, the prediction itself may also serve as a signal of the incumbent's optimistic nature. If voters value optimism among elected officials, as the evidence provided in the previous section suggests, expressing an expectation of a good outcome—irrespective of what actually happens—may indicate to voters that the politician has a positive disposition. Third, if the outcome is positive, it may provide voters with a signal of the incumbent's competence or at least with the reassuring notion that good things happen under the incumbent's watch.

To clarify our claims regarding how these three factors might affect an incumbent's decision to set high expectations, let  $U(\cdot)$  represent an incumbent's utility function. The incumbent can either set a high or low expectation for performance:  $H$ ,  $L$ . The incumbent gains utility based on how his actions affect his constituents; as they are better off, his reelection prospects increase. Let  $\alpha$  represent the value to voters of a correct prediction made by the incumbent,  $\beta$  represent the value from the incumbent setting an optimistic expectation, and  $\gamma$  represent the value from a good outcome.<sup>8</sup> Additionally, let  $p$  represent the probability of there being a bad outcome and  $1 - p$  be the probability of there being a

<sup>8</sup>Note that we actually do not need to include  $\gamma$  in the model since it will drop out when we compare the utilities from incumbents who set high versus low expectations as they both will get credit for good outcomes regardless of their predictions.

**FIGURE 1 Relationship Between Media Mentions and Respondent Ratings of Recent Presidential Candidates' Optimism**



*Note:* The x-axis denotes the fraction of media reports in the year preceding the election about the candidate that mention variants of the term “optimism” divided by the number of media reports that mention the candidate and the term “speech.” The media reports on optimism exclude instances where optimism is referenced in the context of an electoral victory. The y-axis denotes survey respondents’ perceived optimism of the candidates; higher ratings correspond to greater levels of perceived candidate optimism. Vertical grey lines represent 95% confidence intervals.

good outcome. An incumbent’s utilities from setting a high and low expectation are, therefore,

$$U(H) = p\beta + (1 - p)(\alpha + \beta + \gamma); \text{ and} \quad (1)$$

$$U(L) = p\alpha + (1 - p)(\gamma). \quad (2)$$

Note that  $U(H)$  can be rewritten as  $\beta + (1 - p)(\alpha + \gamma)$ . A politician who sets high expectations always receives  $\beta$  but only receives  $\alpha + \gamma$  when the outcome is good. The difference in utilities for setting a high versus low expectation is, therefore,

$$U(H) - U(L) = [p\beta + (1 - p)(\alpha + \beta + \gamma)] - [p\alpha + (1 - p)(\gamma)], \quad (3)$$

which simplifies to

$$U(H) - U(L) = \beta + \alpha - 2p\alpha. \quad (4)$$

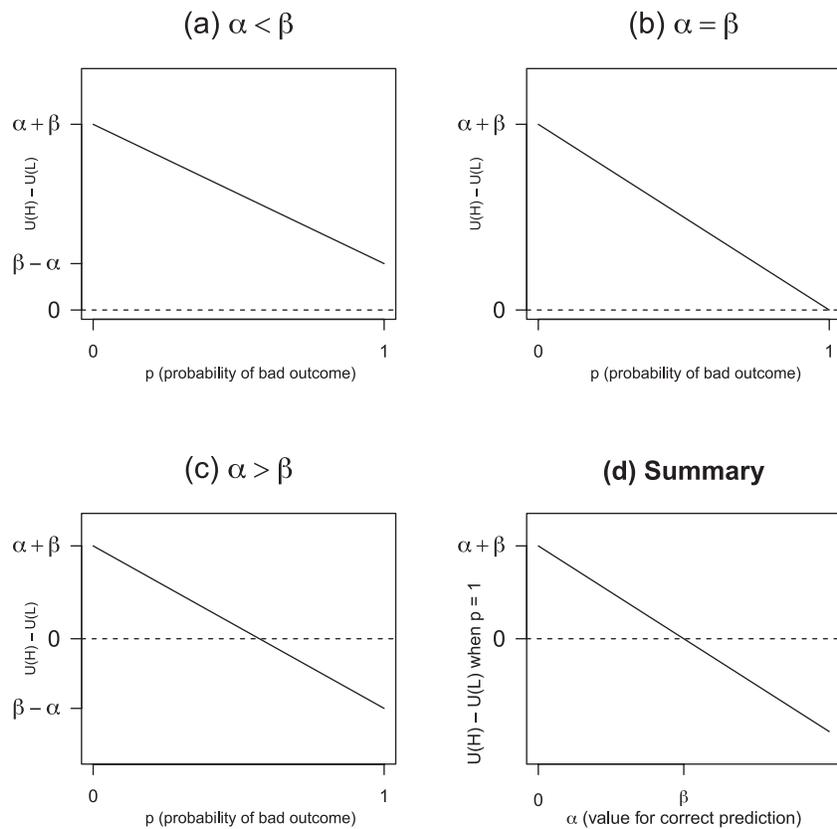
As seen in Equation (4), the relationship between  $p$  and  $U(H) - U(L)$  depends on the relative sizes of  $\alpha$  and  $\beta$ .<sup>9</sup> In Figure 2, we plot the relationship for

<sup>9</sup>Technically speaking, the relationship between  $p$  and  $U(H) - U(L)$  solely depends on  $\alpha$  since the slope of the line represented in Equation (4) is  $-2\alpha$ . However, because  $p$  is bounded between  $[0,1]$  the value of  $\beta$  determines the  $y$ -intercept and therefore the value of  $p$  where  $U(H) = U(L)$ . Below, we refer to this value as the “cut point.”

different relative values of  $\alpha$  and  $\beta$ . As illustrated in Figure 2(a), when  $\alpha < \beta$  (i.e., when the value to a voter of an optimistic politician exceeds the value of making a correct prediction), an incumbent always prefers to set a high expectation regardless of the value of  $p$ . As shown in Figure 2(b), when the two parameters are equal, the incumbent still weakly prefers to set a high expectation in all cases, with the incumbent being indifferent only when a bad outcome is certain ( $p = 1$ ). Conversely, when  $\alpha > \beta$  (i.e., when the value to voters of an incumbent who makes a correct prediction exceeds the value of an optimistic incumbent), there is some point where  $p$  is sufficiently high such that the incumbent prefers to set a low expectation. We refer to this value of  $p$  as the “cut point.” We summarize the theoretical framework in Figure 2(d), where we plot  $U(H) - U(L)$  when  $p = 1$  against  $\alpha$ . When  $\alpha < \beta$ , the incumbent unambiguously prefers to set high expectations ( $U(H) > U(L)$  for all  $p$ ). When  $\alpha > \beta$ , there is some value of  $p$  where  $U(H) < U(L)$ , or some sufficiently high probability of a negative outcome where the incumbent receives less utility from setting a high expectation.

The three aspects of incumbent performance on which our analysis focuses—correct prediction,

FIGURE 2 Theoretical Framework



optimism, and a good outcome—give rise to three different types of domains. The first type is one in which voters perceive the leader as exerting a meaningful impact on the outcome (we refer to these as Type I domains). Most obviously, this is the case when the leader has *practical authority* over the domain (see Christiano [2012] and Hurd [1991] for the distinction between practical and theoretical authority). For example, the U.S. President as Commander in Chief has practical authority over deploying troops. In the second type of domain, the leader does not have practical authority and does not exert direct control over the outcome but is nonetheless perceived as a *theoretical authority* (we refer to these as Type II domains). The theoretical authority can stem from two main sources. It may be due to the leader’s perceived expertise and/or access to information. For example, a prime minister of a small country asked in an interview about the prospects of a peaceful resolution to a civil war in another part of the world is likely to be seen by voters as a theoretical authority on the matter even without having any plausible impact on the security situation in the far-off region. Relatedly, theoretical authority may also be linked to the type of reasoning needed to form a view on the issue at stake.

Even if the leader has no practical authority on the matter, her projection might nonetheless be seen as a demonstration of her judgment process. Voters might therefore make inferences of the form: “If that is how the leader thinks about issue *x*, then she is likely also to make this type of judgment when it comes to issue *z*.” Finally, in the third type of domain the leader is perceived to have no meaningful impact over the outcome and is not perceived as a theoretical authority (we refer to these as Type III domains). For example, a leader asked by a talk show host about the chances of a certain film to win the Academy Award for Best Picture is unlikely to be seen by voters as a theoretical authority on the matter. A failed prediction on such an issue will therefore not reveal much about the leader’s competence nor about his judgment.

It is perhaps worth emphasizing that both concepts of practical and theoretical authority are related to—yet conceptually distinct from—the notion of “clarity of responsibility,” an influential theory in the economic voting literature (Powell and Whitten, 1993). According to this theory, where decision-making authority is concentrated among few political actors (i.e., a setting in which voters can more easily attribute responsibility), economic voting

should be most pronounced. Importantly, the concept of practical authority on which we focus does not necessarily imply that a politician has sole or even dominant control over a domain, but only that her actions can directly influence outcomes within the domain. For example, the national economy can be considered a Type I domain in systems with either high or low clarity of responsibility. Furthermore, the concept of theoretical authority is also quite distinct from clarity of responsibility since even when a politician has little control over outcomes in a domain, she might still be expected to exhibit expertise regardless of the number of political actors that also share responsibility.<sup>10</sup>

We posit that for Type I and Type II domains,  $\alpha > \beta$ . In other words, while voters may value high expectation setting per se on the part of politicians when managing these domains, they derive greater utility from a correct prediction. On the other hand, for Type III domains,  $\alpha < \beta$ . Since few signals about competence or judgement can be extracted from a correct prediction, the value of  $\beta$  is higher than  $\alpha$ . Accordingly, we predict that in Type I and II domains, there will be a cut point ( $\tau$ ) where the probability of a bad outcome is sufficiently large that politicians prefer to set low expectations because they anticipate substantial voter punishment in the case of an incorrect prediction (as in Figure 2c). In contrast, in Type III domains, the cut point will be greater than one, meaning that the politician will always be better off setting high expectations regardless of  $p$  (as in Figure 2a). These predictions are summarized in Table 1. Building on this theoretical framework, we design experiments that exogenously manipulate expectation setting and outcomes as well as cover a diverse set of domains that fall into each of the three categories.

## Experimental Design

### Overview of Studies

We conducted six studies in which we administered our experimental designs via online platforms. As explained below, our basic setup involved 2 x 2

<sup>10</sup>Recent studies have found a weaker relationship between national economic performance and the electoral success of the ruling government in more globalized countries. Researchers interpret this as evidence that voters perceive elected officials as having less of an impact on the the national economy due to the constraining influence of global market forces (Hellwig 2008). Nonetheless, voters could still perceive leaders as being theoretical authorities on local economic issues.

designs in which we manipulated: (1) the expectation set by the elected official (high vs. low) and (2) the eventual outcome and level of performance (good vs. bad). Our studies examined domains for which: (1) the incumbent was plausibly able to exercise a relatively high degree of influence (e.g., the local economy, national security); (2) the incumbent had little control over the outcome but the prediction provides a signal of his judgment (e.g., the global economy, climate change/emissions); and (3) the incumbent had little control over the outcome and expectation setting signals little to nothing about his judgment (e.g., weather, sports outcomes). We examined two main dependent variables: assessments of the leader's judgment and whether the respondent would vote to reelect him.<sup>11</sup> Studies 1–5 were conducted on samples of U.S. adult respondents. Study 6 was conducted on a sample of Spanish adults to provide greater insight on the generalizability of the findings across political contexts.

Study 1 was conducted as part of the 2010 Cooperative Congressional Election Study (CCES), administered by YouGov/Polimetrix to 1,500 respondents over the Internet between October 1, 2010 and November 1, 2010. Study 2 was also conducted by YouGov/Polimetrix between July 15, 2011 and July 20, 2011 using a sample of 1,000 respondents. Study 3 was conducted on a national sample of 1,018 respondents recruited by Survey Sampling International (SSI) between February 16, 2012 and February 18, 2012. Study 4 was conducted using 500 volunteers on Amazon's Mechanical Turk between January 25, 2013 and January 30, 2013. Study 5 was administered by SSI to a U.S. sample of 1,525 respondents between January 21, 2014 and January 21, 2014. Finally, Study 6 was conducted on a sample of 1,086 Spanish adult citizens recruited by Netquest between January 23, 2014 and February 10, 2014.<sup>12</sup> We conducted six studies at different time points using different samples. The consistency of the results demonstrates the robustness of the findings.

<sup>11</sup>We also examined respondent assessments of politicians' leadership abilities as well as general favorability. For space reasons, we have excluded them from the article but the findings are extremely similar to those reported here. These results are reported in online Appendix 8. In Studies 5 and 6, we asked about favorability instead of reelection.

<sup>12</sup>Recent articles published in leading political science journals have leveraged each of the different sample populations employed here: YouGov/Polimetrix (e.g., Jessee 2009), SSI (e.g., Malhotra, Margalit, and Mo 2013), Netquest (e.g., Fernández-Albertos, Kuo, and Balcells 2013), and Mechanical Turk (e.g., Huber, Hill, and Lenz 2012).

TABLE 1 Theoretical Predictions by Domain Type

Examples	Type I	Type II	Type III
	Local Economy National Security	Global Economy Emissions	Weather Sports/Entertainment
Value of optimism ( $\beta$ )	↘	↘	↘
Value of correct prediction ( $\alpha$ )	↘	↘	
Value of positive outcome ( $\gamma$ )	↘		
Relative value of $\beta$ and $\alpha$	$\alpha > \beta$	$\alpha > \beta$	$\alpha < \beta$
Cut point ( $\tau$ )	$\tau < 1$	$\tau < 1$	$\tau > 1$

Via an inductive process, we conducted a series of studies that considered a diverse set of domains, altering different characteristics about the types of signals voters could extract from incumbent performance. To be sure that our results are robust and replicable, we tested each domain type at least twice. In Study 1, we examined the effect of expectation setting (conditional on performance) for domains in which voters would plausibly expect the incumbent to have a great deal of control (Type I domains). In Study 2, we extended our analysis to Type III domains. In Study 3, we replicated both types of domains from Studies 1 and 2, but we also included a Type II domain. Finally, in Study 4, we replicated the analyses from the Type II domain from Study 3, as well as tested an additional Type II domain. The purpose of Studies 5 and 6 was to assess the extent to which the results could be generalized. Study 5 studied all three domain types among U.S. respondents but asked them to evaluate a U.S. governor instead of a prime minister, allowing us to examine whether familiarity with the office affects the results given that some American survey respondents may not be familiar with prime ministers. In Study 6, we administered the same experimental design in Spain to see if the results could be exported to other advanced democracies that have different electoral systems. Accordingly, we first report findings from Studies 1–4 before turning to the issue of generalizability toward the end of the article. In all six studies, we randomized the order in which the various experiments were presented to respondents, thereby accounting for potential spillover effects (Transue, Lee, and Aldrich, 2009).

In assessing the political implications of the various scenarios described in the experiments, one potential concern is that respondents may associate the scenario with an actual politician that they know, potentially introducing pretreatment bias and making it difficult to isolate the effects of the manipulations. To ensure that this type of attribution did not

occur, the object of evaluation in the experiments was explicitly described as either a generic leader of a foreign country (Studies 1–4, 6) or an unidentified governor of a U.S. state (Study 5).

Our goal in these experiments was to utilize scenarios that were analogues to the widely studied cases of retrospective evaluation mentioned above—the stimulus and the Iraq War (Karol and Miguel 2007; Malhotra and Margalit 2010). In those cases, an incumbent was being evaluated prior to reelection, and voters were in a position to consider both predictions and outcomes (perhaps informed by challengers and the media), suggesting that voters often have both pieces of information at their disposal. Of course, this is not always the case. For example, when two challengers are seeking an open seat, neither has a track record with which to benchmark against a prediction, so perhaps in these cases, setting high expectations is unambiguously beneficial. Alternatively, in other instances, incumbents are not forced to make a prediction one way or the other (though the media often compels expectation setting on important issues).

Accordingly, as the first study on expectation setting and retrospective voting, our focus is necessarily narrow in scope. For example, we do not consider: (1) situations where voters learn about expectations and outcomes at different points in time (though, the media and challengers often couple them during elections); (2) how voters react to multiple predictions/outcomes by the same political leader; (3) how partisanship conditions the effects; (4) how challengers affect the expectations of voters regarding the incumbent (e.g., claiming that unemployment will rise if the incumbent wins reelection); and (5) whether the results differ between executives and legislators. The objective, as with any model of politics, is to abstract away certain details to assess the causal effects of key features. Clearly, there are multiple avenues for future research to build upon the results presented here.

## Type I Domains

We first assessed how expectation setting affected evaluations of government performance in domains in which the leader would reasonably be expected to exercise some degree of control, or domains where leaders have practical authority. To that end, we conducted the Local Economy Experiment and the National Security Experiment, both of which were conducted as part of Study 1 and then replicated in Study 3.

**The Local Economy Experiment.** In The Local Economy Experiment, we presented respondents with the following preface and then asked them to evaluate the prime minister:

In a televised speech to the nation several years ago, a prime minister of a foreign country predicted that his country's economy would [*pick up/slow down*] the following year. [*Indeed, in line with that prediction/ However, in contrast to that prediction*], in the 12 months that followed that speech the country's economy performed very [*well/poorly*] on every major economic indicator.

The factors that were manipulated are bolded and in italics for presentational purposes (the respondents did not see them in this format). We manipulate two features of the preface. First, we manipulate whether the prime minister set an optimistic expectation for the country's local economy ("pick up" vs. "slow down"). Second, we manipulate whether the ultimate outcome of the economic performance was that it performed "well" or "poorly." Further, we also mentioned whether the outcome was in line or not in line with the prime minister's prediction. Hence, expectation setting (*E*) can take on the values of {high, low}, and outcomes (*O*) can take on the values of {good, bad}. We discuss how we analyze the experimental data below.

**The National Security Experiment.** In the National Security Experiment, we presented respondents with the following preface and then asked them to evaluate the prime minister, again manipulating *E* and *O*:

In a televised speech to the nation several years ago, a prime minister of a foreign country predicted that entering negotiations with its hostile neighboring state would [*greatly improve/severely harm*] his own country's security situation. [*Indeed/However*], in the 12 months that followed the opening of negotiations, the country's security situation vis-a-vis that neighboring state has gotten [*better/worse*].

Note that in the National Security Experiment, the leader is described as having additional control over

the outcome compared to the Local Economy Experiment. The leader can decide whether to enter into the negotiations. Accordingly, the theory would predict that the penalty to making an incorrect prediction would be even higher in this case.

## Type II Domains

We next assessed how expectation setting conditions evaluations of government performance in domains in which the leader could not be reasonably expected to have meaningful influence over the outcome, but for which voters could learn about the leader's judgment from his prediction—in other words, domains where leaders have theoretical authority but not practical authority. To that end, we conducted the Global Economy Experiment (Studies 3 and 4) and the Emissions Experiment (Study 4).<sup>13</sup>

**The Global Economy Experiment.** In the Global Economy Experiment, we presented respondents with the following preface and then asked them to evaluate the prime minister, again manipulating *E* and *O*:

In a televised speech to the nation several years ago, a prime minister of a foreign country predicted that the global economy would [*pick up/slow down*] the following year. [*Indeed, in line with that prediction/ However, in contrast to that prediction*], in the 12 months that followed that speech the global economy performed very [*well/poorly*] on every major economic indicator.

Unlike a country's local economy, a prime minister can hardly be seen as having a significant impact on the overall performance of the global economy. Therefore, people who are benchmarking properly should consider the global economy to be a domain in which the leader does not have meaningful influence (Kayser and Peress 2012; Leigh 2009). Nonetheless, understanding the international economy is an important skill of a national leader, and making faulty predictions may send a negative signal to voters regarding the leader's judgment.

**The Emissions Experiment.** The second Type II domain we asked about dealt with emissions. As with the global economy, small nations do not substantially affect the global level of pollution. However, small countries are affected by emissions and the possibility of climate change, and voters may think

<sup>13</sup>Participants were randomly presented with only a single question pertaining to the economy, either about the local economy or about the global economy. This was done to minimize the possibility of the questions leading respondents to focus on the difference in the politician's influence over the outcome in question.

that leaders should be knowledgeable about environmental issues. The experiment specifically noted China, a major contributor to greenhouse gasses. We explicitly mentioned that the country was small so that it was salient to respondents that the prime minister could not unilaterally influence the level of global emissions.

A prime minister of a small country gave a speech on global pollution and the harmful effects it was having on the country. The prime minister predicted that China would take steps to [*reduce/increase*] its emissions, thereby [*combating/exacerbating*] global pollution. [*Indeed, in line with that prediction/However, in contrast to that prediction*], China [*reduced its emissions by investing in renewable energies/increased its emissions by expanding industrial growth*].

### Type III Domains

We next assessed how expectation setting affects evaluations of government performance in domains in which the leader could not be reasonably expected to have any influence over the outcome and where voters could not extract any signal about the leader's judgement from the expectations she set. In other words, these are domains in which politicians have neither practical nor theoretical authority. To that end, we conducted the Weather Experiment and the World Cup Experiment. Studies 2 and 3 included both experiments.

**The Weather Experiment.** In the Weather Experiment, we presented respondents with the following blurb and then asked them to evaluate the prime minister, again manipulating *E* and *O*:

In a press conference, a prime minister of a foreign country presided over the opening of a new amusement park. The prime minister predicted that [*good weather that summer would draw many tourists to the country/bad weather that summer would keep many tourists away from the country*]. [*Indeed/However*], [*good weather that summer caused tourism to dramatically rise, boosting the country's economy/bad weather that summer caused tourism to severely drop, harming the country's economy*].

Several recent articles have explored how weather affects incumbents' reelection prospects (e.g., Achen and Bartels 2002; Gasper and Reeves 2011; Healy and Malhotra 2010). However, voters may hold politicians accountable for response to or mitigation of the bad weather and not the natural phenomenon itself. Hence, the vignette was written in a way that made it clear that the prime minister could not be responsible for mitigating any damages arising from the bad

weather and that the weather's impact on tourism is a matter of pure luck. Further, we emphasized in the vignette that bad weather would have substantial effects on the country's economy.

**The World Cup Experiment.** In the World Cup Experiment, we presented respondents with the following preface and then asked them to evaluate the prime minister, again manipulating *E* and *O*:

Prior to the 2010 Soccer World Cup, a prime minister of a foreign country was interviewed about his country's prospects in the tournament. The prime minister predicted that his country's team would [*advance to the latter rounds of the tournament/not advance very far in the tournament*]. [*Indeed/However*], a month later the country's team performed [*very well and reached the latter rounds of the tournament/poorly and got knocked out in the first round*].

This experiment draws on work by Healy, Malhotra, and Mo (2010), who contend that sports outcomes are events irrelevant to incumbent performance that officials cannot be reasonably expected to control and therefore domains over which leaders would be expected to have neither practical nor theoretical authority. It is also possible that the World Cup activates feelings of patriotism.

To assess if individuals categorized the domains as we did, in Study 4, we asked respondents prior to the experiments to report on a 5-point scale "How much expertise should political leaders have in matters related to [DOMAIN]?" (response options: "a great deal"; "a lot"; "a moderate amount"; "a little"; and "none at all"). We asked about the six domains included as part of the experiments ("the economy of their countries"; "their country's national security"; "the global economy"; "global pollution and the environment"; "local weather patterns"; and "their country's sports teams"). Respondents said that political leaders should have the most expertise in Type I domains, followed closely by Type II domains. For the four Type I and II domains, the mean rating exceeds 4 ("a lot") in three cases and for all four exceeds 3.9. Conversely, for the two Type III domains in which no judgment signal can be extracted, the mean respondent ratings were below 2.3. Detailed results can be found in online Appendix 4.

### Dependent Variables

Across the four main studies, we examined two main dependent variables: judgment (Studies 1, 2, and 4) and support for reelection (Studies 2, 3, and 4). We randomly rotated response options. We recoded all

dependent variables to lie between 0 (lowest possible value) and 1 (highest possible value) so that we can easily interpret a regression coefficient as a 100 $\beta$  percentage point change in the dependent variable caused by the treatment.

*Judgment.* “Imagine you were a citizen of that country. What effect, if any, would this set of events have on your confidence in the Prime Minister’s judgment?” (response options: “greatly increase my confidence”; “somewhat increase my confidence”; “no effect on my confidence”; “somewhat decrease my confidence”; and “greatly decrease my confidence”).

*Reelection.* “Imagine you were a citizen of the country and the Prime Minister was seeking reelection. What effect, if any, would the events described above have on your support for the Prime Minister’s re-election bid? (response options: “greatly increase my support”; “somewhat increase my support”; “no effect on my support”; “somewhat decrease my support”; and “greatly decrease my support”).

We also examined two additional dependent variables: leadership (Studies 1 and 4) and favorability (Studies 3 and 4). For space considerations, we report these results in online Appendix 8. The results were extremely similar to those found for judgement and reelection.

## Regression Model

To analyze the experiments, we estimated the following regression model via ordinary least squares (OLS)<sup>14</sup> regression:

$$Y_i = \alpha + \beta_1 B_i + \beta_2 E_i + \beta_3 (B_i \times E_i) + \varepsilon_i, \quad (5)$$

where  $i$  indexes respondents,  $Y_i$  represents one of the two dependent variables,  $B_i$  is a dummy variable coded “1” if the respondent was assigned to the “bad outcome” condition (and “0” if the respondent was assigned to the “good outcome” condition),  $E_i$  is a dummy variable coded “1” if the respondent was assigned to the “high expectation” condition (and “0” if the respondent was assigned to the “low expectation” condition), and  $\varepsilon_i$  represents stochastic error.

We can interpret the parameter estimates of the model as follows.  $\beta_2$  represents the optimism benefit/penalty when the outcome is good, and  $\beta_2 + \beta_3$  represents the optimism benefit/penalty when the

outcome is bad.<sup>15</sup> When the outcome is good, it seems unambiguous that people would value high expectations either as a sign of optimism or because they value competent prediction. Hence, we are able to uniquely identify the optimism benefit/penalty by looking at bad outcomes. In these cases, we can assess if people prefer leaders who set high expectations even when outcomes are bad ( $\beta_2 + \beta_3 > 0$ ), whether the negative effect of bad outcomes are heightened by high expectations ( $\beta_2 + \beta_3 < 0$ ), or whether expectations are irrelevant ( $\beta_2 + \beta_3 = 0$ ).  $\beta_3$  can therefore be interpreted as the “slope coefficient” of how increasing the probability of a bad outcome from 0% likelihood to 100% likelihood affects the optimism benefit that accrues to the leader. Hence, the optimism benefit/penalty in expectation is  $\beta_2 + p\beta_3$  where  $p$  is the probability of a bad outcome occurring. An additional quantity of interest is how likely the bad outcome needs to be in order to make the leader indifferent between setting high and low expectations (i.e., the value of  $p$  such that  $\beta_2 + p\beta_3 = 0$ ). We estimate this likelihood via the following “cut point”  $\tau$ :

$$\tau = -\frac{\beta_2}{\beta_3}. \quad (6)$$

If  $\tau$  is close to or exceeds 1, it means that voters value optimism highly in leaders because even if the likelihood of the bad outcome is close to certain, it is still beneficial for leaders to set high expectations.<sup>16</sup> Note that  $\tau$  is conceptually equivalent to the cutpoint presented in the theoretical section above. Hence, our experimental design allows us to estimate the theoretical parameter of interest.

A standard model of retrospective voting only includes  $B_i$  in Equation (5), positing that voters judge politicians based on outcomes. The innovation in this analysis is to consider whether expectations moderate the relationship between outcomes and performance

<sup>14</sup>For ease of interpretation and the presentation of results, we estimated all regressions via ordinary least squares. We also estimated all regressions via ordered logistic regression and obtained essentially similar results.

<sup>15</sup>The 2 x 2 design means that there are four experimental groups. The mean value of the dependent variable (and associated confidence interval) by group is presented in online Appendix 6. Although each respondent provided multiple evaluations, the design is not conducive to a within-subjects analysis because no respondent evaluated the same domain twice.

<sup>16</sup>We must be careful in interpreting  $\tau$  as a probability since the model does not constrain it to lie between 0 and 1. The best way to think of the estimate of  $\tau$  is in terms of a linear probability model. Estimates of  $\tau$  that significantly exceed 1 can be interpreted as follows: no matter how likely an outcome is to be negative, it is always in a politician’s interest to set high expectations.

assessments, necessitating the inclusion of  $E_i$  as well as the interaction between  $E_i$  and  $B_i$ .

## Testable Hypotheses

Based on the regression model, we test a series of hypotheses based on the parameter estimates. For exposition, we subscript parameters with roman numerals indicating the domain types described above.

Test #1: Does setting high expectations provide a benefit when outcomes are good? We can answer this question by testing the null hypothesis that  $\beta_2 = 0$ .

Test #2: Does setting high expectations provide a penalty when outcomes are bad? We can answer this question by testing the null hypothesis that  $\beta_2 + \beta_3 = 0$ .

Test #3: Is the cutpoint  $\tau < 1$ ? In other words, is there some probability of a bad outcome occurring where individuals prefer leaders who set low expectations (i.e.,  $\alpha > \beta$ )? Or is  $\tau > 1$ —is it unambiguously beneficial for a politician to be optimistic (i.e.,  $\alpha < \beta$ )?

Test #4: Is the cutpoint lower for Type I and II domains compared to Type III domains? We can answer this question by comparing  $\tau_I$  to  $\tau_{III}$  and  $\tau_{II}$  to  $\tau_{III}$ .<sup>17</sup>

## Results

We first examine domains in which people perceive elected officials as having a substantial degree of control over the outcome (Type I domains). We present the relevant parameter estimates from the regression model (with associated standard errors) in Table 2. The raw regression model output which produces the estimates for this and all other domains is presented in online Appendix 5. The first thing to point out is that all the estimates are highly significant in statistical terms; consequently, we focus on the substantive interpretation of the results. Additionally, the results are highly consistent across domain, study, and dependent variable, underscoring the robustness of the results. When outcomes are good, leaders experience a benefit from setting high expectations. The coefficient estimates indicate that these leaders are viewed between 23 and 30 percent-

age points more positively across the various dependent variables (Test #1). Conversely, when outcomes are bad, leaders who set high expectations are viewed between 15 and 26 percentage points more negatively than leaders who set low expectations (Test #2). The final column of Table 1 shows that the negative slope coefficients associated with bad outcomes range from 45 to 51%. Lastly, the estimated cut points suggest that bad outcomes need to have about a 50% chance of occurring before it is disadvantageous for a leader to set high expectations. In each case, we reject the null that  $\tau = 1$  (Test #3).

We next examine results from the domains in which leaders are not perceived to have control over the outcomes but for which expectation setting provides a signal about the leader's judgment (Type II Domains). Parameter estimates are presented in Table 3. In the Global Economy and Emissions experiments, leaders experience significant benefits to setting high expectations when outcomes are good in the range of 26–46 percentage points, and significant optimism penalties when outcomes are bad in the range of 26–38 percentage points. This produces cut points of about 50%, similar to those found for the Type I domains. In each case, we reject the null that  $\tau = 1$ .

Finally, we explore domains in which people could not reasonably expect leaders to have control over (or responsibility for) the outcomes and also for which no signal about judgment or competence could be extracted (Type III domains). Parameter estimates are presented in Table 4. While there is general stability across studies and dependent variables, the results are not exactly the same across experiments, so we will examine them one by one. For the Weather Experiment, we find that the optimism benefit when the outcome is good is between 9 and 11 percentage points and that when the outcome is bad, there is no benefit or penalty (the parameter estimates are all close to zero and statistically insignificant). Additionally, all of the cut points are close to 1 (and indeed cannot be statistically distinguished from 1). This means that even when a bad outcome is certain, leaders are better off setting high expectations.

As with the Weather Experiment, in the World Cup experiment leaders accrued substantively large and significant benefits from setting high expectations when outcomes were good, ranging between 15 and 28 percentage points. Leaders also experienced small optimism benefits when outcomes were bad, ranging between 2 and 5 percentage points (statistically

<sup>17</sup>When testing coefficients from different models, for simplicity we assume there is zero covariance between the coefficient estimates. Because we would most likely expect there to be a positive covariance, our tests of the differences in the parameter estimates are conservative. We prefer this more conservative test over attempting to estimate the covariance via various clustering techniques.

TABLE 2 Optimism Benefit/Penalty in Type I Domains

Domain	Dependent Variable	Optimism Benefit/Penalty			
		Good Outcome ( $\beta_2$ )	Bad Outcome ( $\beta_2 + \beta_3$ )	Cut point ( $\tau = -\frac{\beta_2}{\beta_3}$ )	Slope ( $\beta_3$ )
Local economy	Judgment <sup>†</sup>	.27 (.02)	-.23 (.02)	.54 (.02)	-.51 (.02)
Local economy	Reelection <sup>‡</sup>	.30 (.03)	-.15 (.03)	.68 (.05)	-.45 (.04)
National security	Judgment <sup>†</sup>	.25 (.02)	-.23 (.02)	.52 (.03)	-.48 (.03)
National security	Reelection <sup>‡</sup>	.23 (.02)	-.26 (.02)	.47 (.03)	-.49 (.03)

Note: Standard errors in parentheses. <sup>†</sup>Study 1; <sup>‡</sup>Study 3.

significant in the case of the reelection variable). This again produced cut points that were statistically indistinguishable from 1. Setting high expectations seemed to produce slightly greater benefits for leaders in the World Cup scenarios. As we speculated earlier, this could be because a sporting event of this nature also incorporates feelings of patriotism, suggesting that leaders who do not have high expectations about their team could also be viewed as insufficiently patriotic.

The cut points for the Type I and II domains are significantly smaller than the cut points for the Type III domains, consistent with our theoretical expectations (Test #4). As shown in online Appendix 7, 27 of the 28 direct comparisons of cut points between Type I/II and Type III domains reveal a statistically

significant difference. The one insignificant case barely exceeds the  $p < .10$  threshold. Conversely, the cut points from the Type I and Type II domains are much more similar to one another, with significant differences appearing in a minority of cases.

We summarize the results in Figure 3. As shown in the top row of the figure, for the Type I domains, the cut points are around 0.5, indicating that setting high expectations becomes a liability once the probability of a bad outcome exceeds 50%. The figure also highlights the consistency of the results across domains, samples, and dependent variables. Note that the actual data look very similar to our theoretical prediction in Figure 2(c), consistent with our expectation that  $\alpha > \beta$  for Type I domains. In the middle row of Figure 3, we see similar results for Type II

TABLE 3 Optimism Benefit/Penalty in Type II Domains

Domain	Dependent Variable	Optimism Benefit/Penalty			
		Good Outcome ( $\beta_2$ )	Bad Outcome ( $\beta_2 + \beta_3$ )	Cut point ( $\tau = -\frac{\beta_2}{\beta_3}$ )	Slope ( $\beta_3$ )
Global economy	Reelection <sup>†</sup>	.26 (.03)	-.26 (.03)	.50 (.04)	-.52 (.04)
Global economy	Judgment <sup>‡</sup>	.46 (.03)	-.38 (.02)	.55 (.02)	-.84 (.04)
Global economy	Reelection <sup>‡</sup>	.38 (.03)	-.35 (.03)	.52 (.03)	-.73 (.03)
Emissions	Judgment <sup>‡</sup>	.31 (.03)	-.25 (.03)	.55 (.03)	-.56 (.04)
Emissions	Reelection <sup>‡</sup>	.26 (.03)	-.20 (.03)	.56 (.04)	-.47 (.04)

Note: Standard errors in parentheses. <sup>†</sup>Study 3; <sup>‡</sup>Study 4.

TABLE 4 Optimism Benefit/Penalty in Type III Domains

Domain	Dependent Variable	Optimism Benefit/Penalty			
		Good Outcome ( $\beta_2$ )	Bad Outcome ( $\beta_2 + \beta_3$ )	Cut point ( $\tau = -\frac{\beta_2}{\beta_3}$ )	Slope ( $\beta_3$ )
Bad weather	Reelection <sup>†</sup>	.09 (.01)	.01 (.01)	1.07 (.17)	-.08 (.02)
Bad weather	Judgment <sup>†</sup>	.11 (.02)	.00 (.02)	.98 (.13)	-.12 (.02)
Bad weather	Reelection <sup>‡</sup>	.11 (.02)	-.01 (.02)	.89 (.12)	-.13 (.02)
Sports performance	Reelection <sup>†</sup>	.24 (.02)	.05 (.02)	1.30 (.14)	-.19 (.03)
Sports performance	Judgment <sup>†</sup>	.28 (.02)	.02 (.02)	1.08 (.08)	-.26 (.03)
Sports performance	Reelection <sup>‡</sup>	.15 (.02)	.05 (.02)	1.55 (.30)	-.10 (.03)

Note: Standard errors in parentheses. <sup>†</sup>Study 2; <sup>‡</sup>Study 3.

domains, again consistent with the expectations laid out in Table 1. Conversely, the bottom row of the figure shows that for Type III domains, setting high expectations is the preferable strategy to take regardless of the probability of a bad outcome. Even when the likelihood of a bad outcome is very high, the optimism penalty is not significantly different from zero. Because we failed to reject the null that  $\tau = 1$ , the actual data most closely resembles Figure 2(b), suggesting that  $\alpha = \beta$  for Type III domains or that being optimistic is as important as making a correct prediction.

Our results raise additional questions about the ways expectations condition retrospective voting and the extent to which we can generalize the results presented thus far.

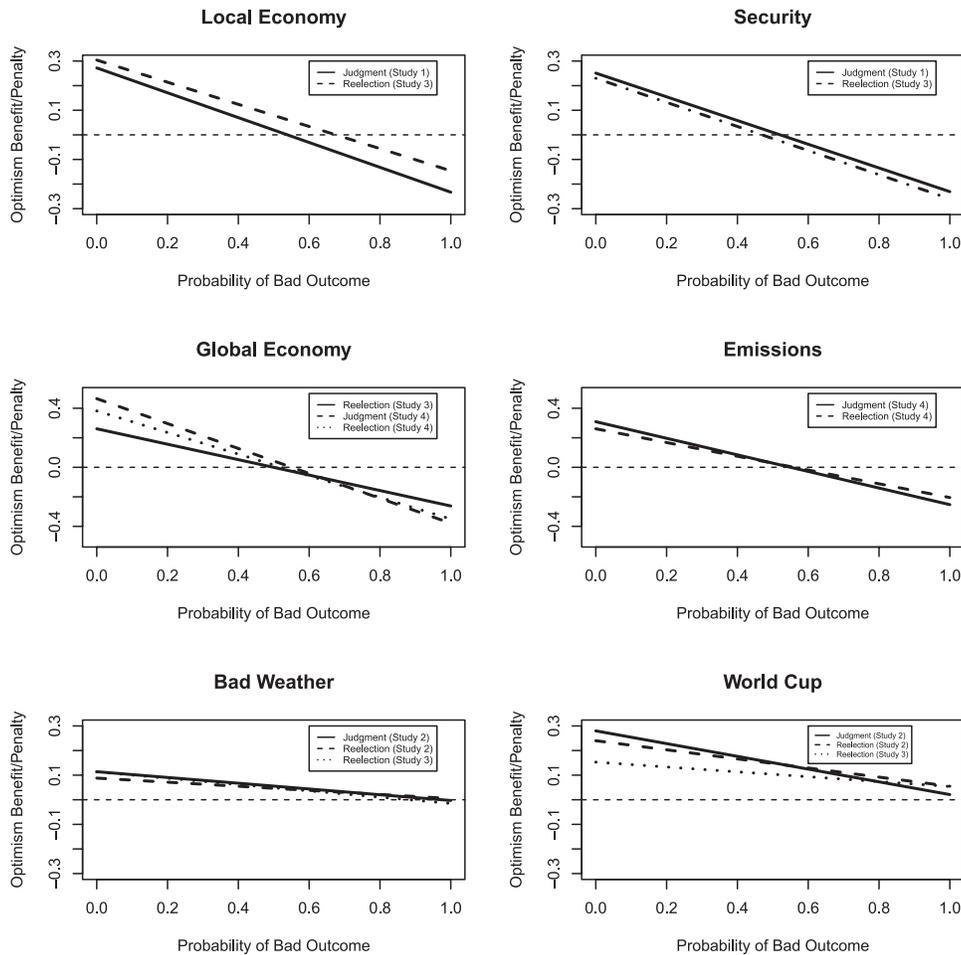
We first explored generalizability across offices. As noted above, in Studies 1–4 we asked U.S. respondents to evaluate a foreign leader so that people would not be primed to think of a particular officeholder but rather about the qualities of political leaders generally. One concern may be that Americans do not fully understand what it means to live in a country with a prime minister as the head of state, and this may affect their responses. To assess if this was the case, we conducted Study 5, where we ask U.S. respondents to evaluate a hypothetical U.S. state governor—an office and political context with which they should be familiar—across three domains: the economy of the governor’s own state (a Type I domain analogous to the Local Economy Experiment above), the U.S. economy as a whole (a Type II

domain analogous to the Global Economy Experiment), and the performance in the NCAA Tournament of a basketball team of a university located in the governor’s state (a Type III domain analogous to the World Cup Experiment). The findings from Study 5 were similar to those in Studies 1–4, suggesting that the results are generalizable across offices and are not sensitive to people’s familiarity with the official position of a given politician. Detailed results can be found in online Appendix 9.

Second, we examined generalizability across countries. The theoretical framework outlined herein is stated generally and is not specific to the U.S. context. To test whether it empirically extends to a different setting, we ran a similar set of experiments in Study 6 in Spain, an established democracy with a different political system. As before, we asked respondents to read a set of vignettes pertaining to a prime minister of a “foreign country,” describing his projections and the eventual outcome in three different domains: the performance of the local economy (Type I), the global economy (Type II), and of the national soccer team in the World Cup (Type III). After receiving this information, respondents assessed the prime minister’s judgement and favorability (the Spanish language translation is presented in online Appendix 11). The findings from the analysis of the Spanish sample are similar to those obtained from the American sample (see online Appendix 9).

Finally, we studied whether politicians reap benefits from lowering expectations and then achieving good results. The  $\alpha$  parameter from Equation (5)

FIGURE 3 Estimated Cut Points by Domain



can provide leverage over this question, but the estimate confounds the benefits voters provide politicians for: (1) being optimistic and (2) making a correct prediction. To address this issue, we included an additional experiment in Study 5 where we presented respondents with three scenarios; in all three, the performance itself exceeds a prediction made by the media in the context of a Type I domain (unemployment). The three scenarios were: (1) a politician sets a low expectation and then exceeds it; (2) a politician sets a high expectation and does not meet it; and (3) a control condition where no expectation is set. The key design feature in this experiment is that in the first and second scenarios, the politician makes an incorrect prediction of equal size, but in one treatment the politician is overly optimistic while in the other he is overly pessimistic. Due to space constraints, details of the study are provided in online Appendix 12. To summarize the results, we find that respondents prefer politicians who set a low expectation and exceed it even when their

expectation is explicitly described as “pessimistic.” These results therefore reinforce the findings described above, namely that in Type I domains politicians are penalized for failing to meet lofty expectations while they are rewarded for exceeding more realistic expectations.

## Discussion

Napoleon Bonaparte famously noted that: “A leader is a dealer in hope” (Bertaut, 1916). Indeed, conventional wisdom holds that politicians must project a healthy dose of optimism if they are to appeal to the voting public. The question that arises is whether setting high expectations by adopting an optimistic stance is a dominant strategy for reelection-minded politicians. The answer is far from obvious. In fact, famed economist and diplomat John Kenneth Galbraith extolled the merits of adopting the exact

opposite strategy: “In public administration good sense would seem to require the public expectation be kept at the lowest possible level in order to minimize eventual disappointment” (Galbraith, 1981, 168). What is the electoral impact of expectation setting by politicians, and what does it imply about optimal vote-seeking strategy?

This article put forward a theoretical framework and leveraged a set of survey experiments to address these questions. Our framework calls for distinguishing between three types of signals that politicians send when setting expectations about a future outcome: regarding their competence, judgment process, and optimistic disposition. We contend that the signals that a prediction encompasses vary systematically across domains. More specifically, domains differ in the extent to which the leader is seen by voters as possessing either practical or theoretical authority. When the leader possesses such authority, signals of competence and/or judgment will be extracted from comparing the prediction to the eventual outcome. In contrast, in domains where the leader possesses neither type of authority, voters solely extract information about the leader’s optimistic disposition. As a result, the rewards or sanctions that voters assign after learning about the accuracy of the leader’s prediction vary systemically across domains. The experimental results presented here provide strong substantiation of this framework using evidence from a broad range of domains and dependent variables, replicating the results over a repeated set of studies across diverse contexts and political settings.

Most directly, these findings contribute to the large body of research on voter retrospection in candidate selection. Notably, this literature focuses on the relationship between outcomes (e.g., the state of the economy during the candidate’s tenure) and voter preferences. However, it ignores the role of expectation setting or the extent to which the prior projections made by politicians affect voters’ ex post assessments of the politician’s performance. Our findings highlight the fact that the expectations set by politicians do in fact moderate voter retrospection and do so in a predictable fashion.

Returning to the examples we presented at the outset, our analysis suggests that President Obama would indeed have been better off electorally had his team not made the (overly) optimistic projections regarding the expected impact of the stimulus on the unemployment rate. Similarly, public sentiment with respect to the Bush administration was likely made worse by its overly optimistic projections regarding

the likely positive response of the Iraqi people to the arrival of American military forces. One might counter that had both Presidents been more sullen (and thus more accurate) in their projections, public support for the policies they sought to advance—economic stimulus and the invasion of Iraq—would have been lower, perhaps even derailing the presidents’ chances of pursuing their respective policies. In other words, their (overly) optimistic projections may have been necessary prerequisites for garnering sufficient political backing to execute the policies.

This conjecture may be true in some instances. Yet it does not invalidate the finding that voters punish mistakenly optimistic projections in domains in which voters attribute to the leader either practical or theoretical authority. An interesting question for future research is whether the high expectations expressed by elected officials regarding certain outcomes are due to short-term needs of garnering political support or whether these projections are in fact a true reflection of the leader’s actual expectations.

Finally, more research is needed on voters’ “default” attribution of responsibility. Given that the leader’s type and degree of authority in a given domain is not always clear cut, our findings raise the question of what the public’s fall-back assumption is. For example, in circumstances where the optimistic leader’s influence on the poor outcome is unclear, do the retrospective assessments of voters reflect a belief that the leader possesses a high degree of influence or not? Exploring the process by which voters gauge a leader’s authority in a given domain would allow us to investigate not only the link between candidate performance and voter attitudes, but also learn more about the role of expectation setting in moderating the electorate’s retrospective assessments.

## References

- Achen, Christopher H., and Larry M. Bartels. 2002. “Blind Retrospection: Electoral Responses to Droughts, Flu, and Shark Attacks.” Paper presented at the Annual Meeting of the American Political Science Association.
- Bertaut, Jules. 1916. *Napoleon in His Own Words*. Chicago: A.C. McClurg.
- Christiano, Tom. 2012. “Authority.” *The Stanford Encyclopedia of Philosophy*.
- Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin. New York: Cambridge University Press, 55–97.

- Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50 (1): 5–25.
- Fernández-Albertos, José, Alexander Kuo, and Laia Balcells. 2013. "Economic Crisis, Globalization, and Partisan Bias: Evidence from Spain." *International Studies Quarterly* 57 (4): 804–16.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.
- Galbraith, John Kenneth. 1981. *A Life in Our Times: Memoirs*. Boston: Houghton Mifflin.
- Gasper, John T., and Andrew Reeves. 2011. "Make it Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters." *American Journal of Political Science* 55 (2): 340–55.
- Harnden, Tony. 2007. "Poll Slump by Bush is Hitting Republicans." *Daily Telegraph* (May 7, 2007).
- Healy, Andrew J., and Neil Malhotra. 2010. "Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence." *Quarterly Journal of Political Science* 5: 193–208.
- Healy, Andrew, and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16 (1): 285–306.
- Healy, Andrew J., Neil Malhotra, and Cecilia H. Mo. 2010. "Irrelevant Events Affect Voters' Evaluations of Government Performance." *Proceedings of the National Academy of Sciences* 107 (29): 12804–09.
- Hellwig, Timothy. 2008. "Globalization, Policy Constraints, and Vote Choice." *Journal of Politics* 70 (04): 1128–41.
- Hennessey, Kathleen, and Christi Parsons. 2012. "President Obama Aims to Accentuate Positive Without Forgetting Negatives." *Los Angeles Times* (February 8, 2012).
- Huber, Gregory A., Seth J. Hill, and Gabriel S. Lenz. 2012. "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106 (4): 720–41.
- Hurd, Heidi M. 1991. "Challenging Authority." *Yale Law Journal* 100 (6): 1611–77.
- Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review* 103 (1): 59–81.
- Karol, David, and Edward Miguel. 2007. "The Electoral Cost of War: Iraq Casualties and the 2004 US Presidential Election." *Journal of Politics* 69 (3): 633–48.
- Kayser, Mark A., and Michael Peress. 2012. "Benchmarking across Borders: Electoral Accountability and the Necessity of Comparison." *American Political Science Review* 106 (3): 661–84.
- Key, V. O. 1966. *The Responsible Electorate*. New York: Vintage Books.
- Kimball, David C., and Samuel C. Patterson. 1997. "Living up to Expectations: Public Attitudes toward Congress." *Journal of Politics* 59: 701–28.
- Kramer, Gerald H. 1971. "Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964." *American Political Science Review* 65 (1): 131–43.
- Leigh, Andrew. 2009. "Does the World Economy Swing National Elections?" *Oxford Bulletin of Economics and Statistics* 71 (2): 163–81.
- Malhotra, Neil, and Yotam Margalit. 2010. "Short-Term Communication Effects or Longstanding Dispositions? The Public's Response to the Financial Crisis of 2008." *Journal of Politics* 72 (03): 852–67.
- Malhotra, Neil, Yotam Margalit, and Cecilia Hyunjung Mo. 2013. "Economic Explanations for Opposition to Immigration: Distinguishing between Prevalence and Conditional Impact." *American Journal of Political Science* 57 (2): 391–410.
- Margalit, Yotam. 2011. "Costly Jobs: Trade-Related Layoffs, Government Compensation, and Voting in U.S. Elections." *American Political Science Review* 105 (01): 166–88.
- Powell, G. Bingham, and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37 (2): 391–414.
- Stolberg, Sheryl Gay. 2011. "And Now, the Cheerleader in Chief." *New York Times* (January 30, 2011).
- Transue, John E., Daniel J. Lee, and John H. Aldrich. 2009. "Treatment Spillover Effects Across Survey Experiments." *Political Analysis* 17 (2): 143–61.
- Tucker, Joshua A. 2006. *Regional Economic Voting: Russia, Poland, Hungary, Slovakia and the Czech Republic, 1990–1999*. New York: Cambridge University Press.
- Waterman, Richard W., Hank C. Jenkins-Smith, and Carol L. Silva. 1999. "The Expectations Gap Thesis: Public Attitudes toward an Incumbent President." *Journal of Politics* 61 (4): 944–66.
- Zullo, Harold M., and Martin E. P. Seligman. 1990. "Pessimistic Rumination Predicts Defeat of Presidential Candidates, 1900 to 1984." *Psychological Inquiry* 1: 52–61.

Neil Malhotra is an Associate Professor of Political Economy in the Graduate School of Business at Stanford University, Stanford, CA 94305.

Yotam Margalit is an Associate Professor in the Department of Political Science at Tel Aviv University, Tel Aviv, Israel.